DOCUMENT RESUME

ED 287 270 EC 200 578

AUTHOR	Haring, Norris
TITLE	Investigating the Problem of Skill Generalization: Literature Review III.
INSTITUTION	Washington Univ., Seattle. Washington Research Organization.
SPONS AGENCY	Department of Education, Washington, DC.
PUB DATE	Apr 87
CONTRACT	300-82-0364
NOTE	74p.; For related documents, see ED 249 737; Literature Review I, ED 249 738; and Literature Review II, ED 265 695.
PUB TYPE	Guides - Non-Classroom Use (055) -- Information Analyses (070)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Educational Trends; Elementary Secondary Education; Evaluation Methods; *Generalization; Meta Analysis; *Self Control; *Severe Disabilities; Skills; *Staff Development; Stimuli; Transfer of Training

ABSTRACT

The third in a series of literature reviews, this monograph presents three articles on skill generalization among individuals with severe disabilities. Kathleen A. Liberty analyzes the results of 15 studies to determine how teaching self-control affected students' performance in training and generalization, "Behavior-Control of Stimulus Events to Facilitate Generalization." Joan E. Kayser and Felix F. Billingsley follow with "Generalization: A Review of Assessment Procedures" which examined 62 studies using generalization assessment procedures from 1970-1975 and 1980-1985. In the final article, "Extending Research Findings: The Role of Staff Development and Evaluation," Valerie Lynch and Frances McCarty consider factors in designing staff development and in conducting comprehensive evaluation plans. (CL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED287270

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

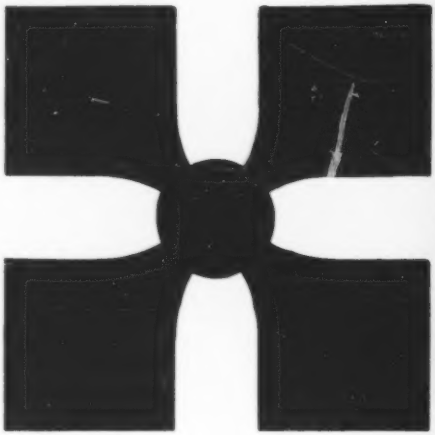
Washington Research Organization

Research in Education of the Severely Handicapped

University of Washington

INVESTIGATING THE PROBLEM OF SKILL GENERALIZATION:

LITERATURE REVIEW III



Norris Haring
Principal Investigator

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Norris G. Haring

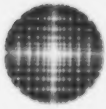
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

EC 200 579

BEST COPY AVAILABLE

INVESTIGATING THE PROBLEM OF SKILL GENERALIZATION:

LITERATURE REVIEW III



Washington Research Organization

Research in Education of the Severely Handicapped

University of Washington
110 Miller Hall, DQ-05
Seattle, WA 98195

INVESTIGATING THE PROBLEM OF SKILL GENERALIZATION:

LITERATURE REVIEW III

Norris Haring, Principal Investigator
Kathleen Liberty, Project Coordinator

Michael Boer, Editor
Jane Kortemeier and Susan Lewis, Typesetters

Many thanks to Valerie Lynch, Claire Phifer, Mary Lynn Shirey, and Zelalem Yilma for their generous assistance in the production of this literature review.

The activity which is the subject of this report was supported in whole or in part by the U.S. Department of Education (Contract No. 300-82-0364). However, the opinions expressed herein do not necessarily reflect the position or policy of the U.S. Department of Education, and no official endorsement by the Department should be inferred.

Printed in the United States of America.

April 1987

Preface

This is the third in a series of literature reviews on the problems of skill generalization. The three components of this review address issues which have developed out of our research and application activities rather than from any preconceived need for conceptual continuity within the literature review series.

Liberty's review of research on self-control, self-monitoring, and self-reinforcement came about as a direct result of her intense interest in the topic. It may ultimately provide a very useful strategy for facilitating generalization. The instructor's need to have effective and cost efficient options for assessing generalization prompted the Kayser and Billingsley review of assessment procedures.

One of the greatest realizations that we have had in our investigations has come as we entered the application phase. It has become clear that the teachers and support staff involved from the participating school districts need additional training in order to apply the recommended intervention procedures. Hence the interest of Lynch and McCarty in studying cost efficiency and durability of training methods for staff development in terms of maintenance of their teaching skills.

Some Trends Since 1977

A summary by White, Leber, and Phifer (1985) of research studies since around 1977 involving a total of 405 subjects shows that substantially more studies targeted functional skills than in the Stokes and Baer (1977) review. In addition, many more studies were with handicapped subjects and were conducted in natural settings, although there still remain a certain number of studies involving skills which are not essential for functioning in natural settings. Since 1977, of the 115 articles having to do with generalization published in 11 journals, 71 (62%) involved severely handicapped students. The two journals that reported the

greatest number of generalization studies were the *Journal of Applied Behavior Analysis* (JABA) and the *Journal of the Association for Persons with Severe Handicaps* (JASH). Forty-eight percent of the articles on generalization which we reviewed came from JABA and 13% came from JASH.

One of the tasks of this institute is to review all of the well controlled and quantified investigations on generalization that have been published since 1977 in terms of whether the critical factors to generalization are likely to contribute to or impede progress toward skill generalization. From analyses of those data, factors and consequating affects have been identified (see table on following page).

Access to multiple settings and/or different teachers and students during the school day does not appear to contribute to skill generalization, unless those variables are identified as part of a particular instructional strategy. Simply providing instruction in a natural setting also does not promote generalization (unless there is only one primary environment); strategies which are designed to improve generalization must be incorporated into instruction. The use of generalization strategies by teachers appears to be more important than the site of instruction in contributing to skill generalization.

Strategies which have been tentatively identified as facilitating generalization include (not necessarily in order of their effectiveness): program natural reinforcers; fade training reinforcers; use natural schedules; use natural consequences; teach self-reinforcement; teach to solicit reinforcement; reinforce generalized behavior; alter contingencies in the generalization situations; vary stimuli using common stimuli, multiple exemplars, or general case approach; increase skill proficiency; fade training stimuli; train in the generalization situation on site; and expand the target skill to increase its function in critical situations.

Factors Contributing to or Impeding Skill Generalization

Factor	Contributing	Impeding
Type of skill instructed	Functional	Not functional
Usefulness in other situations	Useful in many	Useful in one
IEP criteria	Specifies generalization	Does not specify generalization
Level of skill mastery	At or near aim	Acquisition levels
Level of skill fluency	Proficient	Slower than environmental demands
Opportunity to use in other situations	Often	Seldom
Type of instruction	Use strategies for generalization	No strategies for generalization
Consequences in generalization situation	Reinforced for target skill	Not reinforced for target skill
Parents train at home	Happens	Does not happen
Competing behaviors	Controlled or not present	Present

Characteristics of pupil performance which appear to contribute to or impede skill generalization include: level of performance in instruction at time generalizaiton is assessed; relative fluency of target skill and competing skills; type of errors in generalization assessments.

Behaver Control of Stimulus Events

The methods we use in teaching skill acquistion play a critical role in whether or not the S^Ds which cue targeted responses facilitate or impede generalization. As an example, because a specific trainer has become a S^D for the response, his absence from the nontraining setting may result in a lack of appropriate responding. If on the other hand the reinforcement schedules or reinforcers included self-control strategies, the student can play an active role in the mediation of those differences. In a real sense he acts as his own trainer. Since we can't predict what variations future environments will hold, building strategies for the student to use in self-control and decision-making could be an important phase of training.

Assessing Generalization

One area of behavior change in the literature that has been addressed by this report is the review of procedures which are being used to assess generalization. There were a total of 48 articles on assessment of generalization from five journals published from 1980-1985. For comparison we reviewed 14 articles published from 1970-1975. There were 3.4 times more articles which reported data on the assessment of generalization from 1980-1985 than from 1970-1975. This marks a significant increase in the interest of researchers in assessing the generalization of acquired skills. From our review it seems clear that the assessment of generalization, while very time consuming, is important because without the examination of generalization across relevant dimensions in the natural environment, the findings may have limited educational value.

Staff Development

One of the findings of our review is that the sophistication of research on generalization has increased greatly, in particular the research conducted with the severely handicapped.

In applying research findings from the preceding years, we have seen clearly the level of competency required of teachers to apply the processes and procedures that are necessary to ensure that students will generalize skills across settings, across people, and across stimuli. In fact, only two of the teachers involved in the application phase of this study were capable of employing the intervention procedures without extensive inservice training. This observation prompted Lynch and McCarty to conduct a review of the literature on staff development and inservice training.

Our concern with application and replication of the findings in school settings stimulated these questions: Can the new procedures for enhancing generalization be used by public school teachers of the severely handicapped as effectively as the project staff? Can the new procedures produce results similar to the original research findings? Can the new procedures be practical and cost effective enough to ensure widespread application in school districts? In the case of our findings, which involves a long list of strategies found to facilitate generalization, teachers cannot readily determine what strategy to use with which performance problem. Even though a set of decision rules have been developed from the research in this project, in the "application" phase, teachers needed a great deal of assistance in following the rules.

As a result of the application studies and the literature reviews, we have gained valuable information about the staff preparation that should precede the design and implementation of application studies. We probably should have titled this section "I wish I didn't know now what I didn't know then." In any event, the majority of teachers who are currently teaching the severely

handicapped require significantly more training in order to apply strategies which are known to enhance generalization. In addition, the process of deciding which strategy to use in what particular circumstance does involve using rules developed to guide teachers in making that decision with more reliable results. We have seen that teachers can employ these decision rules with specific, systematic, and intensive inservice training.

Norris G. Haring
Principal Investigator
Seattle, 1987

References

Stokes, T. F., & Baer, D. B. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10,* 349-367.

White, O. R., Leber, B. D., & Phifer, C. E. (1985). Training in the natural environment and skill generalization: It doesn't always come naturally. In N. Haring (Principal Investigator), *Investigating the Problem of Skill Generalizing* (3rd ed.). (US Department of Education, Contract No. 300-82-0364). Seattle: University of Washington, College of Education. (ERIC Document Reproduction Service No. ED 265 695)

Table of Contents

Preface v

Table of Contents xi

Liberty, K. A. "Behaver-Control of Stimulus Events to Facilitate Generalization." 1

Kayser, J. E., and Billingsley, F. F. "Generalization: A Review of Assessment Procedure." 17

Lynch, V., and McCarty, F. "Extending Research Findings: The Role of Staff Development and Evaluation." 43

Behaver-Control of Stimulus Events to Facilitate Generalization

Kathleen A. Liberty

Most of the research in the area of stimulus control with severely handicapped people has been directed at the first instances of a response, and how we can manipulate antecedent and consequent events to develop predictable relationships between events and behavior. Research is accumulating which testifies to our success in manipulating stimulus events to promote the acquisition of a broad range of skills by severely handicapped persons–persons generally considered "unteachable" two decades ago.

Our success has brought us new challenges. The very strategies which we use to promote acquisition may interfere with generalization. By using verbal prompts, we may be making it difficult for the student to act when there are no prompts. We can avoid this by fading the prompts, models, demonstrations, and cues we use in instruction. Our use of high density reinforcement during acquisition may also impede generalization. We can gradually reduce our schedule of reinforcement, and also eliminate reinforcers which don't occur in other settings. We should introduce a broad range of stimulus events into training as well, since providing only a few exemplars also causes problems. Research into these and other strategies has been the focus of many of our efforts at solving problems in generalization.

Most of the generalization strategies suggested so far have involved changing how stimulus events are manipulated or presented to the student. An alternative is to change the controller. Instead of control by trainers or teachers, the behaver is taught to control events that may influence generalization. The shift is reflected in the term "self-control." Awkward as this is to use (because when one speaks of "self"-control one is not actually

referring to oneself), it also carries some cognitive and connotative baggage; so the term self-management has come to be used as well. In either case, when that term is used, we are identifying the behaver as the manipulator of antecedent and/or consequent events which may (or may not) have a functional effect on her/his own responding.

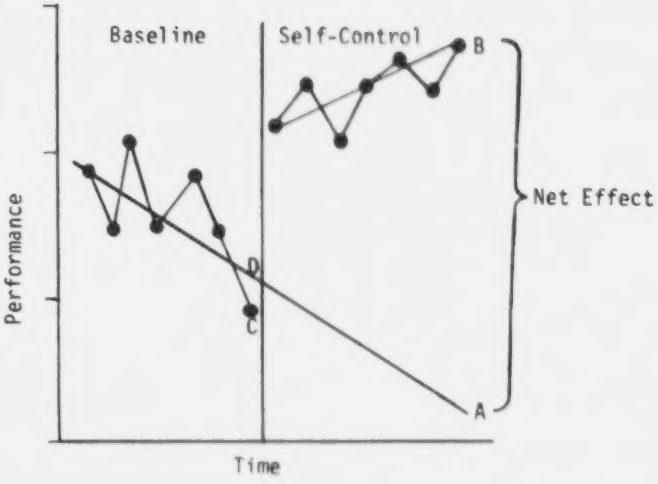
Although many people have touted the "promise" of behaver-control of stimuli, very few research studies have actually investigated generalized responding by behavers who have been taught to control stimulus events. This review analyzes the results of 15 investigations in order to determine how teaching self-control affected students' performance in training and generalization.

We first analyzed seven studies[1] involving 9 subjects. In these studies, the purpose of teaching self-control was to influence behavior directly in the training setting. We also analyzed eight studies[2] involving 16 subjects, in which self-control was taught in order to influence generalization.

In each study, the overall impact of the intervention was calculated by determining the product of the net effect and the median variability. Net effects of teaching self-control were calculated by comparing actual performance at the conclusion of self-management training with performance predicted if baseline conditions had continued during that period, according to the formula: larger divided by smaller (Kazdin, 1976; White, 1971a, 1974; see Figure 1-1). The net effect encompasses changes in both level and trend, and provides a measure of the magnitude of the average effect of the intervention.

The relative variability of performance must also be considered in estimating impact in order to eliminate effects which are actually encompassed by normal variability in performance. In this review, performance variability was calculated for each value in baseline, and the median variability was used to represent the average amount of change predicted by current performance variability.

Figure 1-1
How net effects, variability, and overall impact were assessed



A = Performance predicted by split-middle trend (Kazdin, 1976; White, 1971b, 1972, 1974) if baseline performance had continued without intervention to time at which actual self-control phase ended.

B = Performance at end of self-control phase, calculated at end of split-middle trend.

C = Actual performance.

D = Performance summarized by split-middle trend.

Net Effect = Divide larger of A and B values by the smaller and determine direction of effect.

Variability = Divide larger of C and D values by the smaller for each performance value in baseline.

Overall Impact = Net effect divided by median baseline variability.

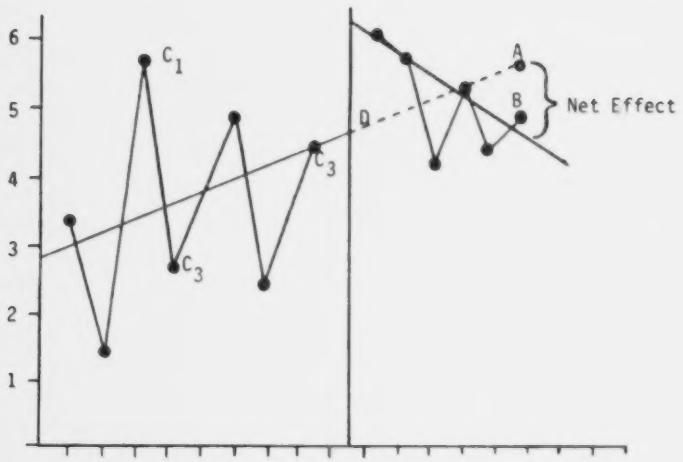
When the magnitude of the net effect is smaller than the magnitude of the daily bounce in baseline, the magnitude of the overall impact is less than 1.0. In these cases, the amount of change during intervention is within the student's normal behavior range prior to intervention (Figure 1-2), and thus the overall impact is probably insignificant. Overall impact was calculated by dividing the net effect, representing changes in both level and trend of performance, by the median baseline variability, representing the relative amount of change predicted prior to intervention as part of the student's normal performance.

Table 1-1 lists the net effect, median baseline variability, and the overall impact on performance in training situations of teaching subjects self-control strategies. Of the 34 performances analyzed, one declined, eight of the changes were within the subject's normal variability of performance, and 25 performances improved. By response class, inappropriate behavior showed the greatest impact; however, expressive communication was the only category in which everyone's performance improved, or showed no change.

Training performance improved for 73.5% of all subjects as a result of self-control training (Figure 1-3). However, only 50% of the severely handicapped subjects' performances improved, as compared to 90% of the other subjects. The magnitude of improvement ranged from 1.1 to 32 times greater than what was predicted from baseline levels.

For 20 of the 34 cases, performance was also assessed in nontraining situations. Generalization to untrained instances, untrained settings, and across time and untrained subjects was included in this sample. Table 1-2 lists the overall impact on performance in generalization. In 16 of 20 instances (80%), generalization improved as a result of self-control training (Figure 1-4). Improvement was shown for 91% of the severely handicapped subjects and 67% of other subjects, just about reversing the proportions whose training performances improved.

Figure 1-2
Insignificant Impact



A = Performance predicted by split-middle trend (Kazdin, 1976; White, 1971b, 1972, 1974) if baseline performance had continued without intervention to time at which actual self-control phase ended.
B = Performance at end of self-control phase, calculated at end of split-middle trend.
C1 = Most variable point.
C2 = Least variable point.
C3 = Middle variable point (i.e., median variability = 1.4).
D = Trend in baseline.

Figure 1-2 shows performance data where net effect is less than median variability during baseline, and thus overall impact is less than 1.0, and not significant.

Table 1-1
Impact of teaching self-control

Article	Subject	Net Effect	Median Baseline Variability	Overall Impact
Work Rate				
1. Bates, Renzaglia, & Clees (1980)[4]	*Subject 1*[1]	(1.1)[2]	1.1[3]	1.0
2. Horner, Lahren, Schwartz, O'Neill, & Hunter (1979)	"Phil"	2.3	1.2	1.9
3. Jackson & Martin (in press)	*Subject 1*	1.1	1.0	1.1
4. Jackson & Martin (in press)	*Subject 2*	1.4	1.0	1.4
5. Jackson & Martin (in press)	*Subject 3*	1.6	1.1	1.5
On Task				
6. Burgio, Whitman, & Johnson (1980)	"Judy" (math)	1.0	1.1	1.0
7. Burgio, Whitman, & Johnson (1980)	"Judy" (phonics)	9.0	1.3	6.8
8. Burgio, Whitman, & Johnson (1980)	"Judy" (printing)	(6.0)	1.2	5.0
9. Fantuzzo, Harrell, & McLeod (1979)	"Ron"	(1.1)	1.4	1.0
10. Burgio, Whitman, & Johnson (1980)	"Angie" (math)	10.0	1.4	7.1

Notes
1. Italicized subjects are severely or profoundly handicapped persons.
2. Values in parentheses indicate that performance worsened. Values without parentheses indicate that performance improved.
3. The higher the value the greater the bounce.
4. Effects of self-administered reinforcement only (excludes effects of changing criterion for reinforcing work rate).

Table 1-1 (continued)

Article	Subject	Net Effect	Median Baseline Variability	Overall Impact
11. Burgio, Whitman, & Johnson (1980)	"Angie" (phonics)	5.0	1.5	3.3
12. Burgio, Whitman, & Johnson (1980)	"Angie" (printing)	3.5	1.2	2.9
13. Horner & Brigham (1979)	Subject A	20.0	1.4	14.3
14. Horner & Brigham (1979)	Subject B	9.0	1.1	8.2
Inappropriate Behavior				
15. Gardner, Clees, & Cole (1983)	Subject 1 (talks to self)	10.0	1.0	10.0
16. Gardner, Clees, & Cole (1983)	Subject 1 (talks to others)	7.0	1.0	7.0
17. Gardner, Cole, Berry, & Nowinski (1983)	*"Roger"*	28.0	1.3	21.5
18. Gardner, Cole, Berry, & Nowinski (1983)	*"Sue"*	40.0	1.6	25.0
19. Rosine & Martin (in press)	*"A"*	1.0	1.2	1.0
20. Rosine & Martin (in press)	*"B"*	(5.0)	3.1	(1.6)
21. Rosine & Martin (in press)	*"C"*	15.0	1.5	10.0
22. Ollendick (1981)	*"David"*	35.0	1.1	32.0

Table 1-1 (continued)

Article	Subject	Net Effect	Median Baseline Variability	Overall Impact
Expressive Communication				
23. Liberty (1984a)	"*Carl*"	1.6	1.5	1.1
24. Liberty (1984a)	"*Cyn*"	1.3	1.0	1.3
25. Liberty (1984a)	"*Lisa*"	1.0	1.2	1.0
26. Liberty (1984b)	"*Joe*"	1.0	1.2	1.0
27. Liberty (1984b)	"*Sam*"	1.0	1.0	1.0
28. Liberty (1983)	"*Shelly*"	1.0	1.2	1.0
29. Harris & Graham (in press)	"Rachel" (action words)	1.4	1.3	1.1
30. Harris & Graham (in press)	"Rachel" (action helpers)	13.0	1.0	13.0
31. Harris & Graham (in press)	"Rachel" (describing words)	4.0	1.1	3.6
32. Harris & Graham (in press)	"Jim" (action words)	1.4	1.3	1.1
33. Harris & Graham (in press)	"Jim" (action helpers)	7.0	1.0	7.0
34. Harris & Graham (in press)	"Jim" (describing words)	2.3	1.3	1.8

Figure 1-3
Impact of self-control on training performance

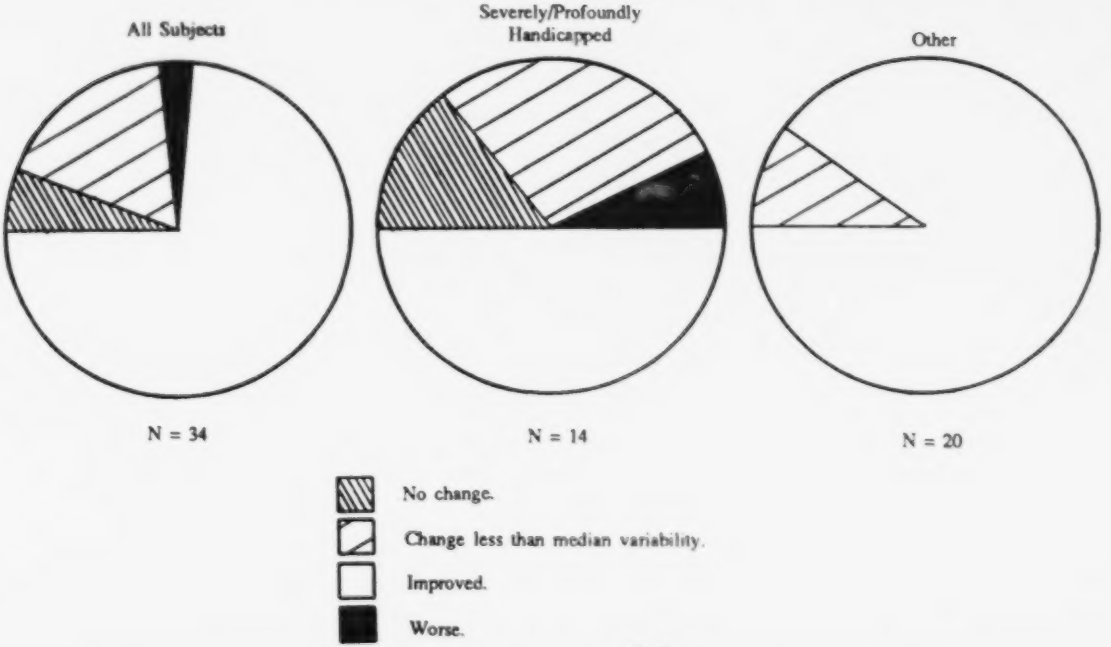


Figure 1-4
Impact of self-control on generalization performance

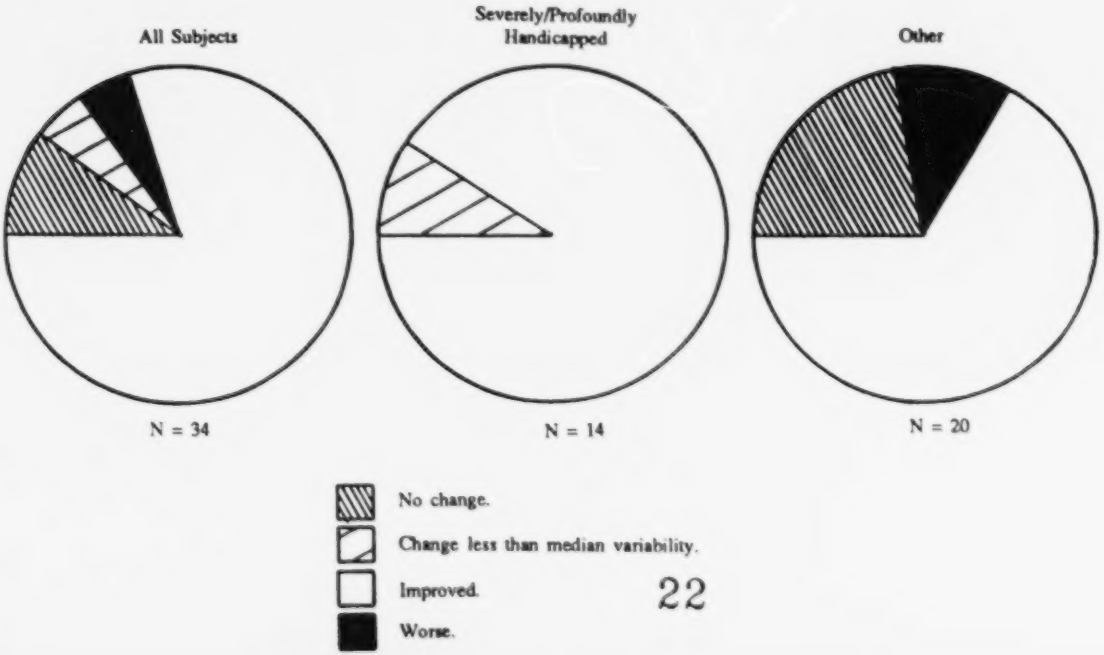


Table 1-2
Impact of self-control on generalization

Article	Subject	Overall Impact: Training	Overall Impact: Generalization	Generalization Dimension
Untrained instances				
1. Liberty (1984a)	"*Carl*"[1]	1.1[2]	2.8	across untrained instances
2. Liberty (1984a)	"*Cyn*"	1.3	1.6	across untrained instances
3. Liberty (1984a)	"*Lisa*"	1.0	1.0	across untrained instances
4. Liberty (1984b)	"*Joe*"	1.0	1.6	across untrained instances
5. Liberty (1984b)	"*Sam*"	1.0	1.5	across untrained instances
6. Liberty (1983)	"*Shelly*"	1.0	1.3	across untrained instances
Untrained Settings				
7. Rosine & Martin (in press)	"*A*"	1.0	15.2	across settings
8. Rosine & Martin (in press)	"*B*"	(1.6)	8.0	across settings
9. Rosine & Martin (in press)	"*C*"	10.0	4.6	across settings
10. Offendick (1981)	"David"	31.2	31.8	across settings

Notes
1. Italicized subjects are severely or profoundly handicapped persons.
2. The figures in this column are repeated from Table 1-1. Values in parentheses indicate that performance worsened. Values without parentheses indicate that performance improved.

Table 1-2 (continued)

Article	Subject	Overall Impact: Training	Overall Impact: Generalization	Generalization Dimension
Across Time				
11. Harris & Graham (in press)	"Rachel" (action words)	1.1	1.0	over time
12. Harris & Graham (in press)	"Rachel" (action helpers)	13.0	15.0	over time
13. Harris & Graham (in press)	"Rachel" (describing words)	3.6	3.8	over time
14. Harris & Graham (in press)	"Jim" (action words)	1.1	1.0	over time
15. Harris & Graham (in press)	"Jim" (action helpers)	7.0	10.0	over time
16. Harris & Graham (in press)	"Jim" (describing words)	1.8	(1.3)	over time
17. Ollendick (1981)	"David"	31.2	31.8	over time (also over settings)
18. Gardner, Cole, Berry, & Nowinski (1983)	*"Roger"*	21.5	21.5	over time
19. Gardner, Cole, Berry, & Nowinski (1983)	*"Sue"*	25.0	25.0	over time
Untrained Subjects				
20. Fantuzzo, Harrell & McLeod (1979)		"Ron" 1.0	"Earl" 2.1	across untrained subject

Not only did self-control instruction improve a greater proportion of generalized performance as compared with training, the magnitude of improvement was greater as well. The median improvement for training was 1.1x while the median for generalization is 2.8x greater than predicted without self-control training.

These results are only suggestive, since data from only a few subjects are available. However, they do suggest that self-control may be effective in facilitating generalization, especially with severely handicapped subjects. These results suggest teaching self-control in order to promote generalization, rather than solely to improve classroom performance.

Generalization may fail to occur for many reasons. For example, because a specific trainer has become an S^D for the response and he is not present in the new situation, or because reinforcement schedules differ across settings. Self-control strategies permit the student to mediate these differences–to act as his own trainer. We can't predict what future environments will bring, so it will require much effort to devise strategies for teachers that will produce generalization to all situations.

We can then require teachers to implement all of these strategies. Self-control offers an alternative to this, too. Instead of trying to predict future situations in which a particular behavior may or may not occur, we should look at who is behaving. The behavior can't occur without a behaver. We could train the behaver in self-control skills so that she, in turn, could cope with future situations which differ from training rather than continuing dependence on trainers to determine environmental events (Dweck, 1975). This is the implicit–and exciting–premise of teaching self-control skills: functional independence.

References

Bates, P., Renzaglia, A., & Clees, T. (1980). Improving the work performance of severely/profoundly retarded young adults: The use of a changing criterion procedure design. *Education and Training of the Mentally Retarded, 15*(2), 95-104.

Burgio, L., Whitman, T., & Johnson, M. (1980). A self-instructional package for increasing attending behavior in educable mentally retarded children. *Journal of Applied Behavior Analysis, 13*, 443-459.

Dweck, C. (1975). The role of expectations and attributions in the alleviation of learned helplessness. *Journal of Personality and Social Psychology, 31*(4), 674-685.

Fantuzzo, J., Harrell, K., & McLeod, M. (1979). Across-subject generalization of attending behavior as a function of self-regulation training. *Child Behavior Therapy, 1*(4), 313-321.

Gardner, W., Clees, T., & Cole, C. (1983). Self-management of disruptive verbal ruminations by a mentally retarded adult. *Applied Research in Mental Retardation, 4*(1), 41-58.

Gardner, W., Cole, C., Berry, D., & Nowinski, J. (1983). Reduction of disruptive behaviors in mentally retarded adults: A self-management approach. *Behavior Modification, 7*, 76-96.

Harris, K., & Graham, S. (in press). Improving learning disabled students' composition skills: A self-control strategy training approach. *Learning Disability Quarterly*.

Horner, R., & Brigham, T. (1979). The effects of self-management procedures on the study behavior of two retarded children. *Education and Training of the Mentally Retarded, 14*(1), 18-24.

Horner, R., Lahren, B., Schwartz, T., O'Neill, C., & Hunter, J. (1979). Dealing with the low production rates of severely retarded workers. *American Association for the Education of the Severely/Profoundly Handicapped Review, 4*(2), 202-212.

Jackson, D., & Martin, G. (in press). Additive effects of components of a self-control package for improving the production of mentally hadnicapped workers. *Journal of Practical Approaches to Developmental Handicaps*.

Kazdin, A. (1976). Statistical analyses for single-case experimental designs. In M. Hersen and D. Barlow (Eds.), *Single case experimental design: Strategies for studying behavior change*. New York: Pergamon Press.

Liberty, K. A. (1983). Self-Monitoring III: Efforts of s-m training on expressive communication: Mediation across settings. In N. Haring (Principle Investigator), *Institute for education of severely handicapped children: Washington Research Organization; Annual Report FY 82-83*. (U.S. Department of Education, Contract No. 300-82-0364). Seattle: University of Washington, College of Education.

Liberty, K. A. (1984a). Effects of a generalization package alone and with self-monitoring on generalized yes/no answers of profoundly handicapped subjects. In N. Haring (Principle Investigator), *Institute for education of severely handicapped children: Washington Research Organization; Annual Report FY 83-84*. (U.S. Department of Education, Contract No. 300-82-0364). Seattle: University of Washington, College of Education.

Liberty, K. A. (1984b). Acquisition and generalization of self-monitoring and self-reinforcement and their effects on question answering. In N. Haring (Principle Investigator), *Institute for education of severely handicapped children: Washington Research Organization; Annual Report FY 83-84*. (U.S. Department of Education, Contract No. 300-82-0364). Seattle: University of Washington, College of Education.

Liberty, K. A., & Michael, L. J. (1985). Teaching retarded students to reinforce their own behavior: A review of process and operation in the current literature. In N. Haring (Principal Investigator), *Investigating the problem of skill generalization* (3rd ed.). (U.S. Department of Education, Contract No. 300-82-0364). Seattle: University of Washington, College of Education.

Ollendick, T. (1981). Self-monitoring and self-administered overcorrection: The modification of nervous tics in children. *Behavior Modification, 5*(1), 75-84.

Rosine, L., & Martin, G. (in press). Self-management training to decrease undesirable behavior of mentally handicapped adults. *Rehabilitation Psychology*.

White, O. (1971a). *A glossary of behavioral terminology*. Champaign, IL: Research Press.

White, O. (1971b) A pragmatic approach to the description of progress in the single case. Unpublished Doctoral Dissertation, College of Education, University of Oregon.

White, O. (1972) *A manual for the calculation and use of the median slope: A method of progress estimation and prediction in the single case*. Regional Resource Center for Handicapped Children, University of Oregon, Eugene, Oregon.

White, O. (1974). *The split-middle: A "quickie" method of trend estimation.* Third Revision. Experimental Education Unit, CDMRC Working Paper. University of Washington, Seattle, Washington.

Footnotes

[1]Studies were those included in an earlier review (see Liberty & Michael, 1985 for criteria for inclusion).

[2]Studies met the following criteria: training and generalization data presented for individual subjects; repeated measures of generalization across phases. These included three studies previously included in UWRO monographs.

Generalization: A Review of Assessment Procedures

Joan E. Kayser
Felix F. Billingsley

The demonstration of generalized treatment effects is becoming a prominent feature in the behavior change literature related to learners with moderate and severe handicaps. However, while there have been advances in methods designed to facilitate generalization (see Stokes & Baer, 1977; Horner, Sprague, & Wilcox, 1982; Liberty, 1985; Warren, Rogers-Warren, Baer, & Guess, 1980), the assessment practices which characterize the investigation of generalization have rarely been subjected to examination (e.g., Kendall, 1981).

In order to identify the procedures currently utilized to assess generalization, all studies which examined generalized performance by learners with moderate to profound handicaps were reviewed in a sample of the experimental literature for a 5-year period, 1980-1985. Forty-eight articles from the following journals were included in the review: *Journal of Applied Behavior Analysis*, *Journal of Autism and Developmental Disorders* (formerly *Journal of Autism and Childhood Schizophrenia*), *Analysis and Intervention in Developmental Disabilities* (initial publication, 1981), and *Journal of the Association for Persons with Severe Handicaps* (formerly *Journal of the Association for the Severely Handicapped* and *AAESPH Review*, initial publication, 1976). In addition, to provide some degree of historical perspective, 14 applicable articles were located in the *Journal of Applied Behavior Analysis* for 1970-1975 and the *Journal of Autism and Childhood Schizophrenia* for 1971-1975 (initial publication, 1971). A bibliography of all the studies surveyed is provided at the end of this review.

For purposes of the review, generalization was defined as "the performance of (previously learned) skills in (previously untaught) new situations—in other school or nonschool settings, with other cues or stimuli, with other individuals, and so on" (Liberty, Haring, & Martin, 1981). "Learners with moderate to profound handicaps" refers to individuals classified as moderately to profoundly retarded, autistic, or multiply handicapped with at least moderate retardation.

Assessment procedures were examined to answer questions in three areas. First, across what dimensions is generalization assessed; that is, is generalization commonly measured across settings, across people, across stimuli, or across a combination of those dimensions? Second, who assesses the learner's performance in nontraining situations, and under what conditions are generalization observations made? Third, when and how often is generalization assessed?

What to Assess

The assessment of generalization has been reported:

1. Across settings (Agosta, Close, Hops, & Rusch, 1980; Dy, Strain, Fullerton, & Stowitschek, 1981; Eason, White & Newsom, 1982; Hill, Wehman, & Horst, 1982; and McGee, Krantz, Mason, & McClannahan, 1983).
2. Across people only (Brady et al., 1984; Breen, Haring, Pitts-Conway, & Gaylord-Ross, 1985; Sternberg, Pegnatore, & Hill, 1983; and Wacker & Berg, 1984a).
3. Across settings and people (Handleman & Harris, 1980; Oliver & Halle, 1982; Strain, 1983; and Warren & Rogers-Warren, 1983).
4. Across stimuli only (Anderson & Spradlin, 1980; Egel, Shafer, & Neef, 1984; Horner & McDonald, 1982; and Hupp & Mervis, 1981).
5. Across settings and stimuli (Coon, Vogelsburg, & Williams, 1981; Schleien, Ash, Kiernan, & Wehman, 1981; and Wacker & Berg, 1984b).

6. Across people and stimuli (Richman, Reiss, Bauman, & Bailey, 1984).

7. Across settings, people, and stimuli (Kleinert & Gast, 1982; Krantz, Zalenski, Hall, Fenske, & McClannahan, 1981; McDonnell, Horner, & Williams, 1984; and van den Pol et al., 1981).

According to the studies surveyed for this review, the most dramatic change in the type of generalization assessed over the past 10 years has been an increase in the measurement of skill generalization across settings. Table 2-1 shows the number of articles for each 5-year period in which skills were measured across each dimension of generalization. In 1970-1975, only 4 of 14 studies (29%) measured generalization across settings, across settings and people, or across settings and stimuli. As indicated in Table 2-1, the most frequently measured dimension of generalization for that time period was across stimuli. For example, Garcia, Guess, and Byrnes (1973) evaluated the generalization of subjects' use of singular and plural declarative sentences to novel items presented by the trainer. In a similar study (Sailor, 1971), generalization of two retarded females' use of plurals in their language was measured across stimuli. In 1980-1985, 33 of 48 investigations (69%) included cross-setting performance in their measurement of generalization. By way of illustration, Tucker and Berry (1980) assessed severely multihandicapped students' ability to put on hearing aids in their residential living units. In a study by Snell (1982), four males with severe mental retardation were taught how to make beds in their classrooms and were assessed for generalization in their dormitories.

Table 2-1
Number of Articles Reporting Generalization Across Each Dimension

Generalization Dimensions	# Studies	
	1970-1975	1980-1985
Across setting only	1	11
Across people only	3	4
Across stimuli only	6	10
Across setting and people	2	12
Across setting and stimuli	1	6
Across people and stimuli	1	1
Across setting, people, and stimuli	0	4
Total =	14	48

Clearly, researchers have become more concerned with ensuring that newly acquired skills are performed outside of training settings. Although not the focus of this review, a comparison of the types of experimental tasks utilized in these studies suggests that the tasks taught from 1980-1985 may have been more conducive to the measurement of cross-setting generalization. In the 1970-1975 literature sample, 11 of 14 (79%) of the experimental tasks were communication-related skills; 2 of 14 (14%) involved decreasing inappropriate behaviors; and, in one study, social interaction skills were instructed. In 1980-1985, the types of behaviors taught were more diverse. Nineteen of 48 (40%) of the experimental tasks were communication-related skills; 9 of 48 (19%) were self-help skills; 7 of 48 (15%) were leisure skills; 6 of 48 (13%) were social interaction skills; 4 of 48 (8%) were vocational skills; and 3 of 48 (6%) involved

decreasing inappropriate behaviors. Even though the highest percentage of experimental tasks taught during each time period involved communication skills, cross-setting generalization was assessed in only 3 of 11 (27%) of the studies in which such skills were taught from 1970-1975, whereas in 1980-1985, 11 of 19 (58%) of the communication skills taught were assessed in nontraining settings. It may be that the increase in the percentage of communication skills which were measured across settings reflects an increase in the number of investigations which focused on functional skills (i.e., skills which might naturally be required in a number of settings rather than experimentally convenient or "developmental" communication skills).

It is interesting that so few studies reported students' performance of skills in nontraining settings, with people not involved in instruction, and with stimuli different from those used in training sessions. In fact, none of the studies in the 1970-1975 sample, and only 4 of 48 (8%) 1980-1985 studies, evaluated generalization across all three dimensions. In a study by Krantz et al. (1981), generalization of verbal descriptors by children with autism was assessed separately and sequentially across new materials, with a teacher not involved in training, and in a nontraining setting. However, van den Pol et al. (1981) assessed moderately retarded students' generalization of restaurant skills simultaneously across settings, stimuli, and people.

The most comprehensive assessment of generalization would be the systematic evaluation of skill performance across each dimension separately (i.e., setting, people, and stimuli), as well as across all three dimensions simultaneously. This could provide information which might indicate possible reasons for unsuccessful generalization or identify otherwise misleading instances of cross-situational responding. For example, in the latter case, if a student's newly learned behavior is controlled by the actions of a specific trainer, and the skill is assessed by that trainer under nontraining setting and stimulus conditions, the data might suggest that the student has successfully generalized the skill. It is possible, however, that the student would not use the skill in situations where the trainer was not present. The systematic evaluation of generalization across the various dimensions might be accomplished by evaluating the student's performance in other settings, but with the same person and

stimuli involved during training; with a person who was not present during training sessions, but in the same setting and with the same stimuli used in training; and with different materials, but with the original trainer present in the training setting. Where generalization does not occur across a particular dimension, then specific generalization facilitation techniques (e.g., general case instruction, programming natural maintaining contingencies, or training loosely) might subsequently be applied. If assessments of generalization are conducted simultaneously across all three dimensions following a systematic analysis related to each condition, it seems likely that the simultaneous assessments will provide a more informative account of generalized performance.

The experimental literature seems to indicate an increasing interest in the use of "functional" tasks in research involving persons with handicaps. Where learners are instructed in the performance of such tasks, consistency with the concepts of social and educational validity require that generalization assessments be conducted within situations in which trained skills would naturally be performed. To illustrate, if toothbrushing is taught in a school environment by a student's teacher, then generalization should be assessed in the student's home, when s/he would naturally engage in the activity, in the presence of the parent or caregiver, and with the materials commonly available in that setting.

Who Should Assess

In our literature sample, the people directly involved in the assessment of generalization included trainers, adults unfamiliar to the learner, and familiar adults and peers not involved in training acting as solicitors and/or observers of the target behavior. In 2 of 14 (14%) of the studies surveyed from 1970-1975, and in 7 of 48 (15%) of the studies surveyed from 1980-1985, spontaneous performance of the target behavior by the student was observed in the generalization setting (i.e. no solicitors were present). For example, in two studies (Foxx, McMorrow, & Mennemeier, 1984; Gaylord-Ross, Haring, Bruce, & Pitts-Conway, 1984), subjects were instructed in social interaction skills and unsolicited interaction with peers was measured in the generalization settings.

In the majority of studies, however, the children were taught behaviors which were evoked by cues from other individuals. The impact of the presence of trainers in the generalization setting is unknown. In 8 of 14 (57%) of the studies reviewed from 1970-1975, and in 27 of 48 (56%) of the studies reviewed for 1980-1985, the trainer solicited the student's behavior when generalization was assessed. In 1970-1975 and 1980-1985, 6 of 14 (43%) and 22 of 48 (46%) of the studies, respectively, reported that the trainers also were responsible for collecting data on subject responses during the generalization probes. For example, in a study by McGee, Krantz, Mason, & McClannahan (1983), receptive labeling of lunch preparation items was solicited and responses recorded during generalization probes by the subjects' house parent, who also provided instruction during the training phase. In another study in which two autistic children were taught "yes" and "no" mands in response to desirable and undesirable food items, subject responses to untrained stimuli were solicited and measured by the original trainers (Hung, 1980). There is some evidence that the implementation of a recording procedure by a trainer may have an effect on the behavior of an "observee" and may result in biased data (Hay, Nelson, & Hay, 1980). In addition, when trainers are present during generalization probes, the studies fail to show that the learner would perform the newly acquired skill in the presence of anyone other than the trainer.

The use of a person unfamiliar to the learner to solicit the target response was reported in only one study from the 1970-1975 sample, and in three studies from the 1980-1985 sample. The presence of unfamiliar observers in the generalization setting was reported in one study from 1970-1975, and in 12 studies from 1980-1985. It is possible that the presence and/or participation of a stranger in the generalization setting may have an impact on the student's performance. For example, the presence of someone unfamiliar to a child in his/her home environment during a generalization probe could potentially distract the child from task requirements. To illustrate, three autistic children's responses to verbal questions were recorded by unfamiliar graduate students in the children's homes in a study by Handleman and Harris (1980). It is possible that the students' highly variable generalization scores may have been due to the presence of strangers in their homes.

A related issue concerns the conditions under which data are collected during generalization probes. In the 1970-1975 literature sample, 10 of 14 articles (71%) reported overt observation, 2 of 14 (14%) reported covert observation, and 2 of 14 (14%) reported both overt and covert observation in the generalization setting. In the 1980-1985 sample, generalization data were collected overtly in 37 of 48 studies (77%), covertly in 3 of 48 studies (6%), and both overtly and covertly in 5 of 48 studies (11%). In 3 of 48 (6%) of the 1980-1985 articles, the manner in which observations were made could not be determined.

Ideally, generalization should be assessed by whomever would naturally be present in the generalization setting. If the learner is expected to perform the skill spontaneously in the generalization setting, then behavioral observations should be conducted covertly, or perhaps overtly and as unobtrusively as possible by someone likely to be present in that situation. For example, in a study by Richman, Reiss, Bauman, and Bailey (1984), institutionalized, mentally retarded females' performance of menstrual care skills was measured by a ward staff member who pretended to be busy "straightening up" in the bathroom.

Frequently children require verbal or physical cues from others in order to perform a newly learned skill. In instances where the behavior is solicited in the generalization setting, the person who would naturally provide cues should solicit the child's behavior. To illustrate, the generalization of manual sign use by profoundly retarded adults in response to verbal instructions was solicited by the teacher in the classroom setting and by a staff member on the ward in a study by Duker and Morsink (1984). When behavior is solicited, the preferable method of collecting generalization data would be covert observation. However, in situations where covert observation is impossible or impractical, data collection by the solicitor may be preferable to overt observation by the trainer or an unfamiliar adult.

When to Assess

Guidelines for determining when and how often to assess for generalization have not been provided in the literature. The articles surveyed for this review indicated that the number of generalization probes ranged from a single, post-training probe session (Eason, White, & Newsom, 1982; Kleinert & Gast, 1982) to multiple probe sessions conducted before, during, and after training (Agosta, Close, Hops, & Rusch, 1980; Brady et al., 1984; Coon, Vogelsberg, & Williams, 1981; Foxx, McMorrow, & Mennemeier, 1984; Stainback, Stainback, Wehman, & Spangiers, 1983; Storey, Bates, & Hanson, 1984; van den Pol et al., 1981). The scheduling of probes in relation to the training phase was diverse for both 5-year periods. As can be seen in Table 2-2, the most dramatic change over the past decade was the relatively large increase in generalization probes conducted after training only (7% for 1970-1975 versus 25% for 1980-1985). It is surprising that generalization probes conducted after training only were as frequently reported as probes conducted before, during, and after training during the 1980-1985 period. The problem with assessing cross-situational performance only after training has occurred is, of course, that no evidence is available to indicate that the learner did not perform the skill in the generalization situation(s) prior to training. In the absence of such evidence, generalization cannot legitimately be claimed.

In the 1980-1985 literature sample, wide variation was noted in the number of generalization probes conducted before, during, and following training. Many investigations contained information that would permit the number of probes employed to be determined only as a range. Therefore, probe ranges were used as a basis for analysis. Where a precise number of probes could be determined, that number was included as both the low and high end of the range. The range across studies for the number of

Table 2-2
Generalization Probe Scheduling

Scheduling of Probes	# Studies	
	1970-1975	1980-1985
Before, during, and after training	3	12
Before and after training	1	7
Before and during training	4	11
During training only	3	2
After training only	1	12
During and after training	2	1
Not reported	0	3
Total =	14	48

probes conducted before, during, and after training, and the medians for the low and high scores of the ranges, are presented in Table 2-3.

The data regarding median number of generalization probes indicate that, on the average, the greatest number was conducted during training, and the smallest number of probes was conducted following training. It is important to determine that skills or parts of skills are generalizing as they are being learned, but it is also of importance to investigate the maintenance of skills in generalization situations following training. Therefore, although post-training generalization probes may extend the length of time required for the completion of investigations, it is recommended that, where possible, at least some generalization probes be conducted over a period of several weeks following training.

The number of generalization probes required during and following instruction may be determined in part by the need to demonstrate the student's consistent performance in the generalization setting. Multiple, consecutive probes may be necessary in cases where the learner must perform the new skill with a high degree of reliability. To illustrate, the generalization of independent street-crossing skills would necessitate multiple, consecutive demonstrations of successful generalization before the student would be allowed to cross streets unobserved.

Table 2-3
Ranges and Low-High Medians of Generalization Probes, 1980-1985 Investigations

Point(s) of Administration	Before Training		During Training		After Training		# of Studies
	Range	Medians	Range	Medians	Range	Medians	
Before, during & after training	1-45	3-7	1-39	6-17	1-20	3-3	12
Before & after training	1-37	3-6	0		2-6	3-4	6*
Before & during training	1-24	4-10	1-124	6-16	0		11
During training only	0		3-12	6-9	0		2
After training only	0		0		1-40	2-3	12
During & after training	0		3-4	-	1	-	1
Not reported							3

*Seven studies reported pre- and post-training probes, but in one study, the number of probes was not reported.

Conclusion

A review of the generalization assessment procedures utilized in a sample of studies from 1970-1975 and 1980-1985 indicated that there has been a recent increase in the measurement of generalization across settings. This increase may be due to the provision of training in skills within a variety of functional task areas that may be required by learners in diverse situations.

There were no dramatic differences found between the two time periods in terms of who assessed generalization, or whether the observations of the learner's behavior in the generalization setting were overt or covert. The fact that trainers were frequently used as solicitors and observers of the target behavior during both 5-year periods suggests that researchers often failed to consider the possible effects of using trainers in the generalization setting.

Over the past decade, a relative increase was noted in the measurement of generalization following training only. Given that data obtained only from post-training assessments provide weak support for assertions of generalized performance, the observed increase in such measurement is alarming.

Comprehensive assessments of generalization may be expensive in terms of the time and participants required, but without the examination of generalization across relevant dimensions in the natural environment, research may result in findings of limited educational utility. It was, therefore, suggested that skill generalization be assessed across settings, across people, and across stimuli, both separately and simultaneously. The person or persons chosen to solicit and/or observe the learner's behavior should be whomever would naturally be present in the generalization setting. Ideally, the recording of subject responses would be accomplished in a covert or unobtrusive manner. In addition, generalization probes should be conducted before, during, and after training to ensure that the target behavior has generalized and is being maintained in nontraining situations. Finally, the number of probes necessary may be determined by the need to demonstrate consistent performance in the generalization setting.

There are still many questions concerning the measurement of generalization that warrant further investigation. How many

examples of generalization across settings, across individuals, and across stimuli are sufficient to demonstrate the learner's successful use of a new skill? It has been suggested that generalization be assessed across all stimulus and response variations which will be encountered in a defined "universe" (Horner, Sprague, & Wilcox, 1982). For some tasks, this is a relatively simple procedure; for other tasks, however, assessment across all possible variations may require considerable time and effort.

The practical logistics of assessing generalization in the natural environment deserve additional attention. Since covert observation of the learner is frequently very difficult or impossible, unobstrusive ways of measuring target behaviors are needed. Other issues related to the observation of skills in the generalization setting include the effects of using the same observers across multiple probes and the results of using the trainer as an observer, even in cases where the observation is covert. In addition, the procedural and scoring reliability of persons who are not trained as solicitors and/or observers may be questionable. Recent research indicated that some parents conducting generalization probes in home settings were unreliable in scoring responses, in administering procedures, or both (Kayser, Billingsley, & Neel, 1986).

The scheduling of generalization probes during the training phase could have an effect on student performance in the generalization setting. Research is needed on the effects of conducting probes before versus after training sessions.

Finally, although not a topic of this review, it should be noted that the issue of reinforcement in the generalization setting has not been resolved in the experimental literature. If students have not been introduced to natural maintaining contingencies, should the reinforcement provided in training also be provided in nontraining situations? If so, does the provision of such reinforcement constitute training, rather than generalization, conditions?

References

Agosta, J. M., Close, D. W., Hops, H., & Rusch, F. R. (1980). Treatment of self-injurious behavior through overcorrection procedures. *Journal of the Association for the Severely Handicapped, 5*(1), 5-12.

Anderson, S. R., & Spradlin, J. E. (1980). The generalized effects of productive labeling training involving common object classes. *Journal of the Association for the Severely Handicapped, 5*(2), 143-157.

Brady, M. P., Shores, R. E., Gunter, P., McEvoy, M. A., Fox, J., & White, C. (1984). Generalization of an adolescent's social interaction behavior via multiple peers in a classroom setting. *Journal of the Association for Persons with Severe Handicaps, 9*, 278-286.

Breen, C., Haring, T., Pitts-Conway, V., & Gaylord-Ross, R.(1985). The training and generalization of social interaction during breaktime at two job sites in the natural environment. *Journal of the Association for Persons with Severe Handicaps, 10*, 41-50.

Coon, M. E., Vogelsberg, R. T., & Williams, W. (1981). Effects of classroom public transportation instruction on generalization to the natural environment. *Journal of the Association for the Severely Handicapped, 6*(2), 46-53.

Duker, P. C., & Morsink, H. (1984). Acquisition and cross-setting generalization of manual signs with severely retarded individuals. *Journal of Applied Behavior Analysis, 17*, 93-103.

Dy, E. B., Strain, P. S., Fullerton, A., & Stowitschek, J. (1981). Training institutionalized, elderly mentally retarded persons as intervention agents for socially isolated peers. *Analysis and Intervention in Developmental Disabilities, 1*, 199-215.

Eason, L. J., White, M. J., & Newsom, C. (1982). Generalized reduction of self-stimulatory behavior: An effect of teaching appropriate play to autistic children. *Analysis and Intervention in Developmental Disabilities, 2*, 157-169.

Egel, A., Shafer, M. S., & Neef, N. A. (1984). Receptive acquisition and generalization of prepositional responding in autistic children: A comparison of two procedures. *Analysis and Intervention in Developmental Disabilities, 4*, 285-298.

Foxx, R. M., McMorrow, M. J., & Mennemeier, M. (1984). Teaching social/vocational skills to retarded adults with a modified table game:

An analysis of generalization. *Journal of Applied Behavior Analysis, 17*, 343-351.

Garcia, E., Guess, D., & Byrnes, J. (1973). Development of syntax in a retarded girl using procedures of imitation, reinforcement and modelling. *Journal of Applied Behavior Analysis, 6*, 311-329.

Gaylord-Ross, R. J., Haring, T. G., Breen, C., & Pitts-Conway, V. (1984). The training and generalization of social interaction skills with autistic youth. *Journal of Applied Behavior Analysis, 17*, 229-247.

Handleman, J., & Harris, S. (1980). Generalization from school to home with autistic children. *Journal of Autism and Developmental Disorders, 10*, 323-333.

Hay, L. R., Nelson, R. O., & Hay, W. M. (1980). Methodological problems in the use of participant observers. *Journal of Applied Behavior Analysis, 13*, 501-504.

Hill, J. W., Wehman, P., & Horst, G. (1982). Toward generalization of appropriate leisure and social behavior in severely handicapped youth: Pinball machine use. *Journal of the Association for the Severely Handicapped, 6*(4), 38-44.

Horner, R. H., & McDonald, R. S. (1982). Comparison of single instance and general case instruction in teaching a generalized vocational skill. *Journal of the Association for the Severely Handicapped, 7*(3), 7-20.

Horner, R. H., Sprague, J., & Wilcox, B. (1982). General case programming for community activities. In B. Wilcox and G. T. Bellamy (Eds.), *Design of high school programs for severely handicapped students* (pp. 61-98), Baltimore, MD: Paul H. Brookes.

Hung, D. W. (1980). Training and generalization of yes and no as mands in two autistic children. *Journal of Autism and Developmental Disorders, 10*, 139-152.

Hupp, S., & Mervis, C. B. (1981). Development of generalized concepts by severely handicapped students. *Journal of the Association for the Severely Handicapped, 6*(1), 14-21.

Kayser, J., Billingsley, F. F., & Neel, R. S. (1986). A comparison of in-context and traditional instructional approaches: Total task, single trial versus backward chaining, multiple trials. *Journal of the Association for Persons with Severe Handicaps, 11*(1), 28-38.

Kendall, P. C. (1981). Assessing generalization and the single-subject strategies. *Behavior Modification, 5*, 307-319.

Kleinert, H. L., & Gast, D. L. (1982). Teaching a multihandicapped adult manual signs using a constant time delay procedure. *Journal of the Association for the Severely Handicapped, 6*(4), 25-32.

Krantz, P. J., Zalenski, S., Hall, L. J., Fenske, E. C., & McClannahan, L. E. (1981). Teaching complex language to autistic children. *Analysis and Intervention in Developmental Disabilities, 1*, 259-297.

Liberty, K. A. (1985). Enhancing instruction for maintenance, generalization, and adaptation. In K. C. Lakin and R. H. Bruininks (Eds.), *Strategies for achieving community integration of developmentally disabled citizens* (pp. 29-71). Baltimore: Paul H. Brookes.

Liberty, K. A., Haring, N. G., & Martin, M. M. (1981). Teaching new skills to the severely handicapped. *Journal of the Association for the Severely Handicapped, 6*(1), 5-13.

McDonnell, J. J., Horner, R. H., & Williams, J. A. (1984). Comparison of three strategies for teaching generalized grocery purchasing to high school students with severe handicaps. *Journal of the Association for Persons with Severe Handicaps, 9*, 123-133.

McGee, G. G., Krantz, P. J., Mason, D., & McClannahan, L. E. (1983). A modified incidental teaching procedure for autistic youth: Acquisition and generalization of receptive object labels. *Journal of Applied Behavior Analysis, 16*, 329-338.

Oliver, C. B., & Halle, J. W. (1982). Language training in the everyday environment: Teaching functional sign use to a retarded child. *Journal of the Association for the Severely Handicapped, 7*(3), 50-62.

Richman, G. S., Reiss, M. L., Bauman, K. E., & Bailey, J. S. (1984). Teaching menstrual care to mentally retarded women: Acquisition, generalization and maintenance. *Journal of Applied Behavior Analysis, 17*, 441-451.

Sailor, W. (1971). Reinforcement and generalization of productive plural allomorphs in two retarded children. *Journal of Applied Behavior Analysis, 4*, 305-310.

Schleien, S. J., Ash, T., Kiernen, J., & Wehman, P. (1981). Developing independent cooking skills in a profoundly retarded woman. *Journal of the Association for the Severely Handicapped, 6*(2), 23-29.

Snell, M. E. (1982). Analysis of time delay procedures for teaching daily living skills to retarded adults. *Analysis and Intervention in Developmental Disabilities, 2*, 139-155.

Stainback, S., Stainback, W., Wehman, P., & Spangiers, L. (1983). Acquisition and generalization of physical fitness exercises in three profoundly retarded adults. *Journal of the Association for the Severely Handicapped, 8*(2), 47-55.

Sternberg, L., Pegnatore, L., & Hill, C. (1983). Establishing interactive communication behaviors with profoundly mentally handicapped students. *Journal of the Association for the Severely Handicapped, 8*(2), 39-46.

Stokes, T. F. & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349-367.

Storey, K., Bates, P., & Hanson, H. B. (1984). Acquisition and generalization of coffee purchase skills by adults with severe disabilities. *Journal of the Association for Persons with Severe Handicaps, 9*, 178-185.

Strain, P. S. (1983). Generalization of autistic children's social behavior changes: Effects of developmentally integrated and segregated settings. *Analysis and Intervention in Developmental Disabilities, 3*, 23-33.

Tucker, D. J., & Berry, G. W. (1980). Teaching severely multihandicapped students to put on their hearing aids. *Journal of Applied Behavior Analysis, 13*, 65-75.

van den Pol, R. A., Iwata, B., Ivancic, M. T., Page, T. J., Neef, N. A., & Whitley, P. F. (1981). Teaching the handicapped to eat in public places: Acquisition, generalization and maintenance of restaurant skills. *Journal of Applied Behavior Analysis, 14*, 61-69.

Wacker, D. P., & Berg, W. K. (1984a). Use of peer instruction to train a complex photocopying task to severely retarded adolescents. *Analysis and Intervention in Developmental Disabilities, 4*, 219-233.

Wacker, D. P., & Berg, W. K. (1984b). Training adolescents with severe handicaps to set up job tasks independently using concepts by severely handicapped students. *Journal of the Association for the Severely Handicapped, 6*(1), 14-21.

Warren, S. F., & Rogers-Warren, A. K. (1983). A longitudinal analysis of language generalization among adolescents with severely handicapping conditions. *Journal of the Association for Persons with Severe Handicaps, 8*(4), 18-31.

Warren, S. F., Rogers-Warren, A., Baer, D. M., & Guess, D. (1980). Assessment and facilitation of language generalization. In W. Sailor, B. Wilcox, & L. Brown (Eds.), *Methods of instruction for severely handicapped students* (pp. 227-257). Baltimore: Paul H. Brookes.

Bibliography

Agosta, J. M., Close, D. W., Hops, H., & Rusch, F. R. (1980). Treatment of self-injurious behavior through overcorrection procedures. *Journal of the Association for the Severely Handicapped, 5*(1), 5-12.

Anderson, S. R., & Spradlin, J. E. (1980). The generalized effects of productive labeling training involving common object classes. *Journal of the Association for the Severely Handicapped, 5*(2), 143-157.

Azrin, N., Besalel, V. A., & Wisotzek, I. E. (1982). Treatment of self-injury by a reinforcement plus interruption procedure. *Analysis and Intervention in Developmental Disabilities, 2*, 105-113.

Barton, E. S. (1970). Inappropriate speech in a severely retarded child: A case study in language conditioning and generalization. *Journal of Applied Behavior Analysis, 3*, 299-307.

Brady, M. P., Shores, R. E., Gunter, P., McEvoy, M. A., Fox, J., & White, C. (1984). Generalization of an adolescent's social interaction behavior via multiple peers in a classroom setting. *Journal of the Association for Persons with Severe Handicaps, 9*, 278-286.

Breen, C., Haring, T., Pitts-Conway, V., & Gaylord-Ross, R. (1985). The training and generalization of social interaction during breaktime at two job sites in the natural environment. *Journal of the Association for Persons with Severe Handicaps, 10*, 41-50.

Browning, E. R. (1983). A memory pacer for improving stimulus generalization. *Journal of Autism and Developmental Disorders, 13*, 427-432.

Carr, E. G., & Kologinsky, E. (1983). Acquisition of sign language by autistic children II: Spontaneity and generalization effects. *Journal of Applied Behavior Analysis, 16*, 297-313.

Charlop, M. H., Schreibman, L., & Thibodeau, M. G. (1985). Increasing spontaneous verbal responding in autistic children using a time delay procedure. *Journal of Applied Behavior Analysis, 18*, 155-166.

Coon, M. E., Vogelsberg, R. T., & Williams, W. (1981). Effects of classroom public transportation instruction on generalization to the natural environment. *Journal of the Association for Persons with Severe Handicaps, 6*(2), 46-53.

Correa, V. I., Poulson, C. L., & Salzberg, C. L. (1984). Training and generalization of reach-grasp behavior in blind, retarded children. *Journal of Applied Behavior Analysis, 17*, 57-69.

Corte, H. E., Wolf, M. M., & Locke, B. J. (1971). A comparison of procedures for eliminating self-injurious behavior of retarded adolescents. *Journal of Applied Behavior Analysis, 4*, 201-213.

Duker, P. C., & Morsink, H. (1984). Acquisition and cross-setting generalization of manual signs with severely retarded individuals. *Journal of Applied Behavior Analysis, 17*, 93-103.

Dy, E. B., Strain, P. S., Fullerton, A., & Stowitschek, J. (1981). Training institutionalized, elderly mentally retarded persons as intervention agents for socially isolated peers. *Analysis and Intervention in Developmental Disabilities, 1*, 199-215.

Eason, L. J., White, M. J., & Newsom, C. (1982). Generalized reduction of self-stimulatory behavior: An effect of teaching appropriate play to autistic children. *Analysis and Intervention in Developmental Disabilities, 2*, 157-169.

Egel, A., Shafer, M. S., & Neef, N. A. (1984). Receptive acquisition and generalization of prepositional responding in autistic children: A comparison of two procedures. *Analysis and Intervention in Developmental Disabilities, 4*, 285-298.

Filler, J., & Kasari, C. (1981). Acquisition, maintenance and generalization of parent-taught skills with two severely handicapped infants. *Journal of the Association for Persons with Severe Handicaps, 6*(1), 30-38.

Foxx, R. M., McMorrow, M. J., & Mennemeier, M. (1984). Teaching social/vocational skills to retarded adults with a modified table game: An analysis of generalization. *Journal of Applied Behavior Analysis, 17*, 343-351.

Freagon, S., & Rotatori, A. F. (1982). Comparing natural and artificial environments in training self-care skills to group home residents. *Journal of the Association for the Severely Handicapped, 7*(3), 73-86.

Garcia, E. (1974). The training and generalization of a conversational speech form in nonverbal retardates. *Journal of Applied Behavior Analysis, 7*, 137-150.

Garcia, E., Baer, D. M., & Firestone, I. (1971). The development of generalized imitation within topographically determined boundaries. *Journal of Applied Behavior Analysis, 4*, 101-112.

Garcia, E., Guess, D., & Byrnes, J. (1973). Development of syntax in a retarded girl using procedures of imitation, reinforcement and modelling. *Journal of Applied Behavior Analysis, 6*, 311-329.

Gaylord-Ross, R. J., Haring, T. G., Breen, C., & Pitts-Conway, V. (1984). The training and generalization of social interaction skills with autistic youth. *Journal of Applied Behavior Analysis, 17*, 229-247.

Giangreco, M. F. (1983). Teaching basic photography skills to a severely handicapped young adult using simulated materials. *Journal of the Association for the Severely Handicapped, 8*(1), 43-49.

Handleman, J., & Harris, S. (1980). Generalization from school to home with autistic children. *Journal of Autism and Developmental Disorders, 10*, 323-333.

Higgs, R., Burns, G., & Meunier, G. (1980). Eliminating self-stimulatory vocalizations of a profoundly retarded girl through overcorrection. *Journal of the Association for the Severely Handicapped, 5*(3), 264-284.

Hill, J. W., Wehman, P., & Horst, G. (1982). Toward generalization of appropriate leisure and social behavior in severely handicapped youth: Pinball machine use. *Journal of the Association for the Severely Handicapped, 6*(4), 38-44.

Horner, R. H., & McDonald, R. S. (1982). Comparison of single instance and general case instruction in teaching a generalized vocational skill. *Journal of the Association for the Severely Handicapped, 7*(3), 7-20.

Hung, D. W. (1980). Training and generalization of yes and no as mands in two autistic children. *Journal of Autism and Developmental Disorders, 10*, 139-152.

Hupp, S., & Mervis, C. B. (1981). Development of generalized concepts by severely handicapped students. *Journal of the Association for the Severely Handicapped, 6*(1), 14-21.

Jackson, D. A., & Wallace, F. R. (1974). The modification and generalization of voice loudness in a 15-year-old retarded girl. *Journal of Applied Behavior Analysis, 7*, 461-471.

Jones, F. H., Simmons, J. Q., & Frankel, F. (1974). An extinction procedure for eliminating self-destructive behavior in a 9-year-old autistic girl. *Journal of Autism and Childhood Schizophrenia, 4*, 241-250.

Keogh, D., Faw, G. D., Whitman, T. L., & Reid, D. H. (1984). Enhancing leisure skills in severely retarded adolescents through a self-instructional treatment package. *Analysis and Intervention in Developmental Disabilities, 4*, 333-351.

Kissel, R. C., Johnson, M. R., & Whitman, T. (1980). Training a retarded client's mother and teacher through sequenced instructions to establish self-feeding. *Journal of the Association for the Severely Handicapped, 5*(4), 382-394.

Kleinert, H. L., & Gast, D. L. (1982). Teaching a multihandicapped adult manual signs using a constant time delay procedure. *Journal of the Association for the Severely Handicapped, 6*(4), 25-32.

Krantz, P. J., Zalenski, S., Hall, L. J., Fenske, E. C., & McClannahan, L. E. (1981). Teaching complex language to autistic children. *Analysis and Intervention in Developmental Disabilities, 1*, 259-297.

Lancioni, G. E. (1982). Normal children as tutors to teach social responses to withdrawn mentally retarded schoolmates: Training, maintenance and generalization. *Journal of Applied Behavior Analysis, 15*, 17-39.

Lancioni, G. E. (1983). Using pictorial representations as communication means with low-functioning children. *Journal of Autism and Developmental Disorders, 13*, 87-105.

Lovass, O. I., Schreibman, L., & Koegel, R. L. (1974). A behavior modification approach to the treatment of autistic children. *Journal of Autism and Childhood Schizophrenia, 4*, 111-129.

McDonnell, J. J., Horner, R. H., & Williams, J. A. (1984). Comparison of three strategies for teaching generalized grocery purchasing to high school students with severe handicaps. *Journal of the Association for Persons with Severe Handicaps, 9*, 123-133.

McGee, G. G., Krantz, P. J., Mason, D., & McClannahan, L. E. (1983). A modified incidental teaching procedure for autistic youth: Acquisition and generalization of receptive object labels. *Journal of Applied Behavior Analysis, 16*, 329-338.

Oliver, C. B., & Halle, J. W. (1982). Language training in the everyday environment: Teaching functional sign use to a retarded child. *Journal of the Association for the Severely Handicapped, 7*(3), 50-62.

Richman, G. S., Reiss, M. L., Bauman, K. E., & Bailey, J. S. (1984). Teaching menstrual care to mentally retarded women: Acquisition, generalization and maintenance. *Journal of Applied Behavior Analysis, 17*, 441-451.

Riechle, J., Rogers, N., & Barrett, C. (1984). Establishing pragmatic discrimination among the communicative functions of requesting, rejecting and commenting in an adolescent. *Journal of the Association for Persons with Severe Handicaps, 9*, 31-36.

Sailor, W. (1971). Reinforcement and generalization of productive plural allomorphs in two retarded children. *Journal of Applied Behavior Analysis, 4*, 305-310.

Schleien, S. J., Ash, T., Kiernen, J., & Wehman, P. (1981). Developing independent cooking skills in a profoundly retarded woman. *Journal of the Association for the Severely Handicapped, 6*(2), 23-29.

Snell, M. E. (1982). Analysis of time delay procedures for teaching daily living skills to retarded adults. *Analysis and Intervention in Developmental Disabilities, 2*, 139-155.

Stainback, S., Stainback, W., Wehman, P., & Spangiers, L. (1983). Acquisition and generalization of physical fitness exercises in three profoundly retarded adults. *Journal of the Association for the Severely Handicapped, 8*(2), 47-55.

Sternberg, L., Pegnatore, L., & Hill, C. (1983). Establishing interactive communication behaviors with profoundly mentally handicapped students. *Journal of the Association for the Severely Handicapped, 8*(2), 39-46.

Stokes, T., Baer, D. M., & Jackson, R. L. (1974). Programming the generalization of a greeting response in four retarded children. *Journal of Applied Behavior Analysis, 7*, 599-610.

Storey, K., Bates, P., & Hanson, H. B. (1984). Acquisition and generalization of coffee purchase skills by adults with severe disabilities. *Journal of the Association for Persons with Severe Handicaps, 9*, 178-185.

Strain, P. S. (1983). Generalization of autistic children's social behavior changes: Effects of developmentally integrated and segregated settings. *Analysis and Intervention in Developmental Disabilities, 3*, 23-33.

Tracey, D. A., Briddell, D. W., & Wilson, G. T. (1974). Generalization of verbal conditioning to verbal and nonverbal behavior: Group therapy with chronic psychiatric patients. *Journal of Applied Behavior Analysis, 7*, 391-402.

Tucker, D. J., & Berry, G. W. (1980). Teaching severely multihandicapped students to put on their hearing aids. *Journal of Applied Behavior Analysis, 13*, 65-75.

Twardosz, S., & Baer, D. M. (1973). Training two severely retarded adolescents to ask questions. *Journal of Applied Behavior Analysis, 6*, 655-661.

van den Pol, R. A., Iwata, B., Ivancic, M. T., Page, T. J., Neef, N. A., & Whitley, P. F. (1981). Teaching the handicapped to eat in public places: Acquisition, generalization and maintenance of restaurant skills. *Journal of Applied Behavior Analysis, 14*, 61-69.

Wacker, D. P., & Berg, W. K. (1984). Use of peer instruction to train a complex photocopying task to severely retarded adolescents. *Analysis and Intervention in Developmental Disabilities, 4*, 219-233.

Wacker, D. P., & Berg, W. K. (1984). Training adolescents with severe handicaps to set up job tasks independently using picture prompts. *Analysis and Intervention in Developmental Disabilities, 4*, 353-365.

Warren, S. F., Baxter, D. K., Anderson, S. R., Marshall, A., & Baer, D. M. (1981). Generalization of question asking by severely retarded individuals. *Journal of the Association for the Severely Handicapped, 6*(3), 15-22.

Warren, S. F., & Rogers-Warren, A. D. (1983). A longitudinal analysis of language generalization among adolescents with severely handicapping conditions. *Journal of the Association for Persons with Severe Handicaps, 8*(4), 18-31.

Wheeler, A. J., & Sulzer, B. (1970). Operant training and generalization of a verbal response form in a speech-deficient child. *Journal of Applied Behavior Analysis, 3*, 139-147.

Whitman, T. L., Mercurio, J. R., & Caponigri, V. (1970). Development of social responses in two severely retarded children. *Journal of Applied Behavior Analysis, 3*, 117-122.

Extending Research Findings: The Role of Staff Development and Evaluation

Valerie Lynch
Frances McCarty

In education, the findings of basic and even applied research come from settings devoted to research and with greater structure than the usual public school setting. Promising new educational techniques are frequently disseminated to potential consumers, education professionals, with little understanding of their feasibility, viability, impact, and costs in the public school setting.

Application or impact studies offer researchers an opportunity to further test their findings in school settings in order to determine if new methods or procedures (a) can be used by staff, (b) produce results similar to original research findings, and (c) are cost effective. Two elements characterize application studies. First, the researcher relies on methods typically used to introduce changes or innovations into school systems. This translates into the use of staff development as a means of conveying a new technique to school personnel. Second, the investigator minimizes the intrusiveness of the research on the time commitments and activities of the teacher.

In addition to the usual research concerns, the investigator conducting an application study must attend closely to (a) the design and implementation of a staff development program which successfully conveys to teachers the method or procedure being investigated, and (b) the design and implementation of a comprehensive plan of evaluation which addresses questions of usability, impact, and cost effectiveness of the technique under study. In this paper we attempt to provide information about staff

development and evaluation which will benefit the researcher interested in application studies.

Staff Development

Conducting application studies in the public schools requires that researchers carefully consider how to effect changes in teacher behavior. In other words, researchers must address the question of how to alter the behavior of public school staff so they can successfully implement techniques identified through basic or applied research. An understanding of what is known about staff development can assist the investigator in successfully transferring a promising method or procedure into the hands of teachers.

What is Staff Development?

In education, the term "staff development" is frequently used interchangably with "inservice training" and "inservice education" by authors and practitioners alike. Although some authors would take exception to this practice (Feiman, 1981; Wade, 1984), these terms will be considered synonomous throughout this paper.

In order to plan staff development activities which result in the successful implementation of practices and methods in schools, it is important to have a clear understanding of what staff development is and its relation to the needs of the researcher. Definitions by various authors provide a basis for such an understanding. Four elements are frequently included in these definitions: (a) purpose; (b) approach; (c) beneficiaries; and (d) context.

Purpose. Personal growth of individuals within education, professional growth of school employees, organizational growth or change, and social change have all been cited as purposes of inservice education. To design inservice training and evaluate its impact, it is critical to clearly articulate its purpose (Fox, 1981).

The most commonly identified reason for staff development is the professional growth of educators (Brookfield, 1981; Dillon-

Peterson, 1981; Edlefelt, 1981; Rude, 1978). Preservice training cannot be expected to equip teachers for the myriad of demands and changes that await them during their professional careers. Therefore, continuing professional development must occur for teachers to successfully fill their complex profesional roles and meet unanticipated changes produced through technological, political, and social shifts which influence education. Inservice education designed to produce professional growth must address goals and objectives which are intrinsic to those who participate (Fox, 1981).

Although personal growth of educators has been defined as one purpose of staff development (Brookfield, 1981; Dillon-Peterson, 1981), it cannot be the sole purpose of staff development. Only when compatible with needs for professional growth can personal development be considered a viable reason for inservice education.

Organizational change or growth has recently received much attention as a focus for inservice education in connection with calls for school improvement. Dillon-Peterson (1981) summarizes the relationship between staff development and organizational change by stating: "Staff development and organization development are a gestalt of school improvement; both are necessary for maximum growth and effective change. They are complementary human processes, inextricably complex" (pp. 2-3). If organizational improvement or change is a desired outcome, the goals of staff development will be intrinsic to the organization and may be extrinsic to the individual professional goals of the educators involved (Fox, 1981).

Instances of social change implemented in educational settings, at least in part through staff development, include bilingual education, desegregation, and education for the handicapped. If social policy implementation is desired, then goals for staff development need only reflect the legal goals of society and not necessarily the goals of the organization or individual professionals (Fox, 1981).

Approach. It is generally agreed that staff development is a systematic process to achieve specific changes or purposes

(Brookfield, 1981; Dillon-Peterson, 1981; Harrison, 1980). There are, however, variations in the definition of steps within the process.

Berman and McLaughlin (1978), Edge and Fink (1978), Ehrenberg and Brandt (1976), and Wood, Thompson, and Russell (1981) have all described different approaches to staff development. The common elements in their descriptions are (a) that inservice education is systematic, following an order or sequence, and (b) that the process of inservice education is cyclical—completion of the steps within the process reinitiates the process.

Beneficiaries. The direct beneficiaries of staff development are obviously those individuals who participate in such activities. The ultimate beneficiaries of staff development, however, are students. Benefits are improvement in students' lives (Feiman, 1981), achievement of student outcomes (Ehrenberg & Brandt, 1976), and impact on student achievement (Rude, 1978; Wood & Thompson, 1980).

Context. Staff development does not take place in a vacuum. Its context is defined by both its purposes and beneficiaries. It is the organization (i.e., school system) or a subsystem (e.g., an individual school, department) which provides the backdrop for staff development—whose clients and goals shape inservice education (Brookfield, 1980; Dillon-Peterson, 1980).

Implications. Staff development for most researchers concerned with evaluating and validating educational methods and procedures might well be defined as a systematic process which, in the context of the needs and goals of the school, promotes the professional growth of staff and ultimately results in benefits for students. This definition has several implications for the design of application studies:

1. Staff development for application studies must follow a systematic process.
2. The goals and objectives of staff development must be stated in terms of changes on the part of participating educators—attitudinal, cognitive, and/or performance outcomes.

3. The goals and objectives of staff development should reflect the needs for professional development of those individuals participating.

4. Because the context of staff development is the school system, it is important to ensure that the goals of staff development are not in conflict with the needs and goals of the organization.

5. Evaluation of staff development should include determination of the degree to which professionals change and the impact of such changes on students.

Staff Development Practices

Application studies, by design, place the researcher in the role of spectator. The researcher must sit on the sidelines of the classroom observing whether teachers use the method or procedure to which they've been introduced, the fluency with which teachers use it, and the degree to which students change as a result of the use of the method or procedure. The major "point of control" for the researcher conducting an application study is the introduction of the the method or procedure to teachers—in other words, the inservice training of teachers. In order to design staff development activities which produce desired changes in teachers (e.g., attitudes, learning, and/or performance), the researcher should be acquainted with current best practices in inservice education.

The information presented below relies heavily on three sources which provide summaries of research on staff development in education. Lawrence (1974) reviewed 97 studies and evaluation reports published between 1962 and 1974 to determine the characteristics of successful inservice education programs. In 1980, Harrison reviewed 47 studies of staff development conducted between 1969 and 1979, quantified the results of these studies, and synthesized the findings. This process, known as meta analysis, was also used by Wade (1984) in her analysis of 91 studies published between 1968 and 1983.

Both Harrison (1980) and Wade (1984) use calculations of effect size to describe how dependent variables were affected by independent variables. To determine effect size, the mean difference between treated and control groups is divided by the standard deviation of the control group or some approximation of this measure.

In their meta analyses, Harrison (1980) and Wade (1984) each identified four dependent variable classifications. Wade (p. 190) uses the following levels of evaluation as dependent variables:

1. **Reaction**: Measures of how the participants feel about the staff development activities, usually subjective.
2. **Learning**: Objective and quantitative measures that assess how much a participant has learned as a result of inservice activities.
3. **Behavior**: Objective measures that document whether or not participants change their behavior as a result of a staff development intervention.
4. **Results**: Objectively determining the effects of staff development on students of participating teachers or on the working environment.

Harrison (1980) classified dependent variables as affective, cognitive, performance, and consequence. His definition of cognitive variables matches Wade's (1984) definition of learning level variables, and performance variables match the definition of behavior by Wade. Harrison's definition of consequence variables is similar to Wade's definition of results, but includes only measures of student performance (e.g., tests of cognitive knowledge and attitude scales) and not effects on the working environment. Harrison's definition of affective dependent variables has no analogue in Wade's classification. Affective variables are related to "changes in interest, attitudes, and values, and the development of appreciations" (Harrison, 1984, p. 58) by participants in staff development activities.

Authors vary in the labels they attach to staff development practices and the categories they use to summarize their findings. For the purposes of this paper, we have tried to avoid jargon in labeling practices and used the following categories to organize information about staff development: arrangements, initiation, instructor, participant characteristics, planning and management, and design.

Arrangements. Both on-site, within the school, and off-site inservice training appear to produce professional growth in teachers. However, on-site inservice education produced greater positive effects on teacher attitudinal outcomes (Harrison, 1980; Lawrence, 1974), teacher performance or behavior outcomes (Harrison, 1980; Lawrence, 1974; Wade, 1984), and cognitive or learning outcomes by teaching (Harrison, 1980; Wade, 1984).

Draba (1975) has called for limiting the number of participants in inservice training, while Berman and McLaughlin (1978) have identified a need for a critical mass of participants. Harrison's (1980) and Wade's (1984) findings are inconclusive with respect to the number of participants to include in staff development activities. Although not statistically significant, Wade's data indicate that, in general, larger groups (over 20) produce slightly higher positive effects.

The time at which staff development activities are scheduled does not produce significant differences in effect size. However, when Wade (1984) analyzed the relation of schedules to evaluation measures, she found some statistically significant results. Reaction effects (i.e., the reaction of participants to training activities) were most positively influenced by training on weekends, evenings, or a combination of times. Weekends and combination schedules produced the most positive effects on participant learning. Participant behaviors were most positively affected by weekend training and training during the school day.

Neither the total duration of training in hours nor the length of training over a period of days were found to significantly influence effect size (i.e., measures of effectiveness based on statistics reported in each reviewed study in a meta analysis). Both Harrison (1980) and Wade (1984) found slightly higher effect sizes for short-term training (i.e., six months or less) as

compared to long-term (i.e., over six months). In terms of the number of hours of training, Wade found that longer treatments were associated with a general lessening of effect size.

Initiation. When examining the initiators of staff development activities, both Harrison (1980) and Wade (1984) found that the majority of programs were outside-initiated (i.e., originated by state/federal government, university researcher, etc.) as opposed to initiation within the school by participants, an administrator, or a supervisor. Harrison found a higher effect size for inside-initiated programs, although only four of a total of 65 programs fell into that category. Wade, on the other hand, found that outside-initiated programs (460 cases) produced almost twice the effect size as inside-initiated programs (174 cases). She also found some statistically significant results when analyzing the source of program origination in relation to evaluation measures. The reaction of participants was most favorable when staff development activities were initiated by the state or federal government. Staff development originated by university researchers produced the greatest positive effects on participant behaviors. An administrator or supervisor originating a staff development program produced the greatest positive effects on participant and/or student results, with university researchers producing the next highest positive results.

Instructor. Supervisory staff (12 cases), college personnel (36 cases), and teachers (5 cases) as instructors of staff development activities were found to be more effective than state department staff (1 case), consultants (5 cases), and intermediate school district staff (2 cases) by Harrison (1980). Wade (1984) included the category of self-instruction in her meta analysis and found this to produce the highest positive effect size followed by supervisory staff and college personnel. Teachers (57 cases) as instructors produced only a small positive effect size.

Participant characteristics. A study by the Rand Corporation of federally funded programs aimed at creating innovation or change in educational practices within school districts (Berman & McLaughlin, 1978) concluded that it is easier to implement and continue changes at the elementary level than at the secondary level. This finding is corroborated by Wade (1984). Wade also

examined the effects of combined participant groups of both elementary and secondary personnel. This category produced the greatest effect size and, when analyzed by outcome measures, produced statistically significant results in terms of participant reaction, learning, and behavior. Elementary school groups produced statistically significant effects on participant and/or student results.

Although staff development activities which are voluntary rather than obligatory have been called for (Draba, 1975), the research findings do not support this. Harrison's analysis (1980) resulted in higher positive effects for obligatory participation, while Wade's data (1984) show a higher effect size for voluntary participation in inservice education.

Increased status appears to be the most effective incentive for inservice training, followed by college credit (Wade, 1984). McLaughlin and Berman (1977) found that release time from normal work activities rather than monetary incentives were used by school districts which were more effective in their inservice training and in maintaining change. Wade's analysis also indicates that release time produces greater effects than pay incentives. In terms of participant reactions to staff development activities, the incentives of status, college credit, and release time significantly enhance effect sizes. Behavior effect sizes are significantly influenced by status and college credit as incentives.

Planning and management. Participant involvement in the planning and management of staff development activities has frequently been suggested (Draba, 1975; Hutson, 1981; Wood & Thompson, 1980) and appears to be supported by research. Teacher participation in project decision-making was found to strongly correlate with the effective implementation and continuation of innovative projects in the Rand study (Berman & McLaughlin, 1978). Lawrence (1974) found that school-based inservice education programs conducted by personnel from outside the district were more effective when teachers or administrators were involved as helpers and planners than those programs in which they provided no assistance.

Wade (1984) analyzed the relationship between trainer and participants in staff development activities. According to her

analysis, trainers who assume a superordinate role (i.e., in a position of authority or expertise) in relation to participants will produce a higher, statistically significant mean effect size than if a collegial relationship is assumed. Reaction, learning, and behavior effects were all more positively influenced when the trainer was in a superordinate role. When participant and/or student results are desired, a collegial relationship between the trainer and participants is more effective.

Harrison (1980), Lawrence (1974), and Wade (1984) all found that active roles (i.e., generating ideas, behaviors, materials) for participants of inservice education appear to produce greater positive effects than passive roles (i.e., accepting ideas or behavioral prescriptions). Although her findings were not statistically significant, Wade's results suggest that an active participant role produces larger effect sizes for learning and behavior outcomes, while a passive or receptive role produces larger effect sizes for participant reactions to training.

Wade (1984) found that staff development which has an instructional focus on affective techniques produces significantly greater effect sizes for participant reactions to training, participant learning, and participant and/or student results. Participant behaviors are significantly influenced when the instructional focus of inservice education is on improving general teaching rather than on improvement of a specific subject or affective techniques.

Goals of inservice training programs can be viewed as common (i.e., participants working toward a common end/goal) or individual/personal (i.e., participants working toward different goals based on their individual needs). Both Harrison (1980) and Wade (1984) found that common goals produced a larger total effect size than individual goals, although this finding was not statistically significant.

Design. The designer of inservice education must decide whether activities included in training will be common to all participants, individualized for different teachers, or a combination of these approaches. When common and individualized approaches are compared, the results are conflicting. Lawrence (1974) found that individualized activities were more likely to achieve the objectives of training, while Wade (1984) and Harrison (1980) found that

common activities produced a higher mean effect size. Yet in Wade's meta analysis, common activities produced a much higher effect size on behavioral outcomes than did individualized activities, although not statistically significant.

Harrison was the only author to examine the effects of a combination of approaches. Although caution must be exercised in drawing any conclusions because of the small number of studies (7 cases), combined activities produced a higher effect size than either of the other two approaches.

Numerous instructional methods for staff development are available to the inservice educator. Fortunately, a number of authors have examined the effectiveness of different methods of instruction and some commonality exists in their findings.

Joyce and Showers (1980) reviewed over 200 studies of training methods both alone and in combination, and concluded that effective inservice education results when several or all of the following methods are combined: theory presentation/skill description; modeling/demonstration; practice; feedback; and coaching (i.e., classroom-based, hands-on assistance provided to facilitate the transfer of skills/strategies to the classroom). More specifically, they found that the combination of all five methods produced mastery of new ways of teaching, while only modeling, practice, and feedback were required to produce improvement in or fine tuning of skills.

In a second review of training studies, Joyce and Showers (1983) examined instructional methods which produced horizontal transfer of skills (i.e., situations in which skills can be directly applied from the training setting into the workplace) and vertical transfer of skills (i.e., situations in which skill adaptation must occur in order to successfully apply skills acquired in the training setting to the workplace). The most frequently combined instructional methods to achieve horizontal skill transfer were theory, practice, and feedback. Theory, demonstration, practice, feedback, and coaching were most frequently combined to produce vertical transfer.

Lawrence (1974) found that inservice education programs using demonstration, practice, feedback, and books as instructional

methods were likely to achieve a high degree of success. Observation, demonstration, lecture, and books (with only five cases) produced the highest positive effects of the 13 instructional methods investigated by Harrison (1980). In Wade's (1984) examination of 15 different instructional methods, observation, micro teaching, video/audio, and practice proved significantly more effective than other instructional methods, while lectures, games, discussions, and guided field trips produced significantly lower effect sizes than other methods.

Authors who investigated the effects of practice as an instructional method found it to be very effective. The effectiveness of demonstration as an instructional method is also a consistent finding, if one considers observation as a form of demonstration.

Evaluation

Evaluation as an established field of applied social research has grown rapidly over the past 20 years (Raizen & Rossi, 1981), and the importance of evaluation research in education is widely recognized (Bernstein & Freeman, 1975; Gersten, Carnine, & Williams, 1982; Gersten & Hauser, 1984; Raizen & Rossi, 1981; White, 1984; Williams & Elmore, 1976). The purpose of evaluation research is to measure the effects of a program or intervention; that is, to what degree have the changes intended by the intervention been achieved and to what extent can these changes be ascribed to the intervention (Raizen & Rossi, 1981; Weiss, 1972)?

The term "comprehensive evaluation" refers to studies that include three components: monitoring; impact; and ex post facto cost-benefit or cost-effectiveness analyses (Rossi, Freeman, & Wright, 1979). A comprehensive evaluation provides data to determine whether the intervention was carried out as planned, whether the intervention resulted in changes in the intended direction, and what the intervention costs were in relation to its benefits.

Formative Evaluation: Monitoring Program Implementation and Service Delivery

The literature is replete with admonitions that attention be paid to implementation in program evaluation and the reasons why programs should be monitored. First, monitoring is needed for accountability purposes (e.g., who is getting what and how; Rossi, Freeman, & Wright, 1979). Second, monitoring evaluations are generally prerequisites to effective impact assessments, since the failure of programs is often due to faulty performance or nonimplementation rather than ineffective interventions (Rossi & Wright, 1977). Third, monitoring information may be a supplement to, or the sole basis for, deciding whether to continue programs (Carlo, 1977; Roos, Roos, Nicol, & Johnson, 1978).

Monitoring the delivery of services to evaluate the degree of program implementation is undertaken for a number of purposes. A large proportion of programs that fail to show impact are really failures to deliver the interventions in the manner specified. There are three potential failures: (a) none (or not enough) of the intervention is implemented, (b) the wrong intervention is implemented, or (c) the intervention is unstandardized, uncontrolled, or varies across implementation. In each instance, the need to monitor the delivery of services and identify discrepancies is essential (Rossi, Freeman, & Wright, 1979). An intervention may perform poorly at a given school or site, but without formative evaluation it is unclear whether the performance is due to problems inherent in the intervention or problems in the way the intervention was implemented at that particular site (House, Glass, McLean, & Decker, 1978; Louck & Hall, 1977; Proper, 1980). As a growing body of evidence suggests that educational innovations are rarely, if ever, implemented exactly as planned, there is a need to collect actual implementation data (Gersten & Hauser, 1984).

Impact Evaluation

The extent to which an intervention is used depends on a number of factors. One factor that is critical is evidence of effectiveness; that is, the program outcomes and the conditions under which implementation occurs to produce those outcomes (Wang &

Ellett, 1981). Impact evaluation is the assessment of the extent to which an intervention results in desired changes in the target population. Questions that need to be asked in an impact evaluation include: (a) Is the intervention effective in reaching the intended goals, (b) can results be explained by other variables which are not part of the intervention, and (c) has the intervention resulted in unintended effects? For an intervention to have impact, it must result in movement toward desired objectives.

When conducting an impact evaluation, there must be a plan for the collection of data. The data collection plan should allow the investigator to demonstrate that the outcomes that occurred were the result of the intervention, and to reject any competing explanations or confounding effects. Therefore, impact evaluations need to be undertaken as systematically and rigorously as possible in order to document the causal linkages between intervention inputs and program outcomes.

The critical issue in impact evaluation is whether or not a program has produced significantly more of an effect than would have occurred without the intervention. Two prerequisites to an effective impact evaluation are having: (a) goals that are clearly defined so that it is possible to measure goal attainment, and (b) evidence that the intervention is sufficiently implemented. Initiating impact evaluations requires the identification and explication of one or more outcome measures that both reflect the intervention goals and which are sensitive enough to allow measurement of change if the intervention is effective.

Impact evaluation for pupils with severe/profound handicaps. The most common method of impact assessment has been the comparison between large-*n* experimental and control groups (Rossi, Freeman, & Wright, 1979; White, 1984).

For program evaluation in the area of the severely handicapped, group comparison designs are generally not feasible. As White (1984) points out,

> The greatest problem in the application of "large-n" approaches to the evaluation of programs serving handicapped populations lies in the simple fact that the total number of such children in any given program unit is likely

> to be small. In a hypothetical district of 20,000, it is very unlikely that two or three severely/profoundly handicapped individuals could be found who matched each other reasonably well on even the most obvious educationally relevant variables. [That problem] virtually nullifies the possibility of utilizing the vast majority of traditional strategies for the evaluation of programs serving the severely handicapped population.

An alternative to group comparison designs that has been suggested by White and others (Campbell & Stanley, 1966; Fitz-Gibbon & Morris, 1978; Hersen & Barlow, 1976) is single subject time series analysis. A series of measures, usually given at equal intervals before and after the intervention, is called a time series. A series of measures systematically taken before a program starts can actually eliminate the need for a control group. The single subject time series design uses the students in the program as their own control group, what White calls a "perfect match." Richard Jones (1979) advocated using single subject designs for formative evaluation of individual program components and then following up with group designs to evaluate the overall effectiveness of a program in helping groups of students. (For a detailed discussion of the pros and cons of different research evaluation designs, see White, 1984). Whatever the particular research design, it is important for educators to use evaluation to determine how much students learn and whether this learning can be likened to a particular educational approach (Gersten & Hauser, 1984).

Cost Analysis Procedures in Education

As the resources now available to school districts are scarce, decisions on alternative uses of limited resources need to be made. In the field of education, therefore, cost factors are now analyzed when making program decisions. Cost factors in both new educational programs and proposed changes in existing programs involve looking at several categories of cost over time (Haller, 1974; Levin, 1983; Sorensen & Binner, 1979). Research and development costs are those resources required to develop a program sufficiently for introduction into the system. Investment costs are those necessary to implement the program (e.g., special equipment, training, etc.). Finally, there are operating costs, those recurring costs required to operate a program over time.

The first step in deciding on a model for cost analysis is to decide what model for evaluation would be most appropriate. Two models are used predominantly in the social sciences—cost benefit and cost effectiveness approaches (Alkin, 1970; Barkdoll, 1980; Haller, 1979; Sweeny & Blaschke, 1980). In cost benefit analysis, there is an evaluation of alternatives when costs and benefits are measured in monetary terms. It requires deciding the value of such things, and assigning a dollar value to educational outcomes is a subjective process at best (Weinrott, et al., 1983). Cost effectiveness analysis involves evaluation of alternatives according to both their costs and their effects with regard to producing some outcome or set of outcomes (Schnell, et al., 1979). Under cost effectiveness analysis, both the costs and effects of alternatives are taken into account in evaluating programs with similar goals. It is assumed that (a) only programs with similar or identical goals can be compared and (b) a common measure of effectiveness can be used to assess them (Alkin, 1970; Schnelle, et al., 1979; Sorensen & Binner, 1979; Weinrott, Jones, & Howard, 1983).

Summary

For the educational researcher, application studies are a logical and critical extension of basic and applied research. They provide information on the general utility, feasibility, cost effectiveness, and potential adaptations of educational methods and procedures prior to their broad dissemination. To conduct application studies, the researcher must consider how to design: (a) staff development activities which will enable educators in schools to implement the methods and procedures of basic and applied educational research, and (b) a comprehensive evaluation plan which provides information on the degree of implementation, the level of impact, and the relative costs of the technique under study.

References

Alkin, M. C. (1970). Evaluating the cost-effectiveness of instructional programs. In M. C. Wittrock & D. E. Wiley (Eds.), *The evaluation of instruction: Issues and problems* (pp.221-237). New York: Holt, Rhinehart & Winston.

Barkdoll, G. L. (1980). Cost-benefit analysis and other promising utensils. *How to Evaluate Education Programs*, Ch. 39, p.6.

Berman, P., & McLaughlin, M. W. (1978). *Federal programs supporting educational change, Volume VII: Implementing and sustaining innovations*. Santa Monica, CA: Rand Corporation. (ERIC Document Reproduction Service No. ED 159 289)

Bernstein, I. N., & Freeman, H. E. (1975). *Academic and entreprenaurial research*. New York: Russell Sage.

Brookfield, J. (1981). *Staff development: A systematic process*. Seattle, WA: Western States Technical Assistance Resource (WESTAR), College of Education, University of Washington.

Campbell, D. T. & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Dillon-Peterson, B. (1981). Staff development/organization development–perspective 1981. In B. Dillon-Peterson (Ed.), *Staff development/organization development* (pp. 1-10). Alexandria, VA: Association for Supervision and Curriculum Development.

Draba, R. E. (1975). Guidelines for viable inservice education. *Journal of Reading, 18*, 368-371.

Edelfelt, R. A. (1981). Six years of progress in inservice education. *Journal of Research and Development in Education, 14*(2), 112-119.

Edge, D., & Fink, A. H. (1978). Taking inservice training into the classroom. *Viewpoints in Teaching and Learning, 54* (4), 113-118.

Ehrenberg, L. M., & Brandt, R. S. (1976). Improved student learning: A necessary goal of staff development. *Educational Leadership, 34*(3), 205-209.

Feiman, S. (1981). Exploring connections between different kinds of educational research and different conceptions of inservice education. Journal of Research and Development in Education, 14(2), 11-21.

Fitz-Gibbon, C. T. & Morris, L. L. (1978). *How to design a program evaluation*. Beverly Hills, CA: Sage.

Fox, G. T., Jr. (1981). Purpose as context for evaluating inservice programs. *Journal of Research and Development in Education, 14*(2), 34-45.

Gersten, R. M., Carnine, D. W., & Williams, P. B. (1982). Measuring implementation of a structured educational model in an urban school district: An observational approach. *Educational Evaluation and Policy Analysis, 4*(1), 67-79.

Gersten, R., & Hauser C. (1984). The case for impact evaluations in special education. *Remedial and Special Education, 5*(2), 16-24.

Haller, E. J. (1974). Cost analysis for educational program evaluation. In W. J. Popham (Ed.), *Evaluation in education: Current applications*. Berkeley, CA: McCutchan.

Harrison, D. (1980). Meta analysis of selected studies of staff development (Doctoral dissertation, University of Florida, 1980). *Dissertation Abstracts International, 41*(09), 2804A.

Hersen, M., & Barlow, D. H. (1976). *Single case experimental designs: Strategies for Studying Behavior Change*

House, E. R., Glass, G. V., McLean, L. D., & Decker, F. W. (1978). No simple answer: Critique of the "follow through" evaluation. *Harvard Educational Review, 48*(2), 128-160.

Hutson, H. M., Jr. (1981). Inservice best practices: The learnings of general education. *Journal of Research and Development in Education, 14*(2), 1-10.

Jones, R. R. (1979). Program evaluation design issues. *Behavioral Assessment, 1*, 51-56.

Joyce, B., & Showers, B. (1980). Improving inservice training: The messages of research. *Educational Leadership, 37* (5), 379-385.

Joyce, B. R., & Showers, B. (1983). *Power in staff development through research on training*. Alexandria, VA: Association for Supervision and Curriculum Development.

Lawrence, G. (1974). *Patterns of effective inservice education: A state of the art summary of research on materials and procedures for changing teacher behaviors in inservice education*. Tallahassee, FL: Florida State Department of Education. (ERIC Document Reproduction Service No. ED 176 424)

Levin, H. M. (1983). *Cost-effectiveness: A primer* (New perspectives in evaluation, Vol. 4). Beverly Hills, CA: Sage.

Loucks, S., & Hall G. E. (1977). A developmental model for determining whether the treatment is actually implemented. *American Educatorial Research Journal, 14*, 264-276.

McLaughlin, M., & Berman, P. (1977). Retooling staff development in a period of retrenchment. *Educational Leadership, 35*(3), 191-194.

Proper, E. C. (1980, September). *Remarks on measurement of implementation of standardized programs*. Paper presented at the American Psychological Association, Montreal, Canada.

Raizen, S. A., & Rossi, P. H. (1981). *Program evaluation in education: When? How? To what ends?* Washington, DC: National Academy Press.

Roos, L. L., Jr., Roos, N. P., Nicol, P., & Johnson, C. (1978). Using administrative data banks for research and evaluation: A case study. In T. D. Cook et al. (Eds.), *Evaluation studies review annual, Volume 3* (pp. 263-270). Beverly Hills, CA: SAGE Publications.

Rossi, P. H. & Wright, S. R. (1977). Evaluation research: An assessment of theory, practice, and politics. *Evaluation Quarterly, 1*(February), 5-52.

Rossi, P. H., Freeman, H. E., & Wright, S. R. (1979). *Evaluation: A systematic approach*. Beverly Hills, CA: Sage.

Rude, C. R. (1978). Trends and priorities in inservice training. *Exceptional Children, 45*, 172-176.

Schnelle, J. F., Kirchner, R. E., Galbaugh, F., Domash, M., Carr, A., & Larson, L. (1979). Program evaluation research: An experimental cost-effectiveness analysis of an armed robbery intervention program. *Journal of Applied Behavior Analysis, 12*(4), 615-623.

Sorenson, J. E., & Binner, P. R. (1979). Cost, cost outcome, cost effectiveness. In G. Landsberg, W. D. Neigher, R. J. Hammer, C. Windle, & J. R. Woy (Eds.) *Evaluation in practice: A sourcebook of program evaluation studies from mental health care systems in the United States*. (U.S. DHEW Pub. No. (ADM) 78-763). Washington, DC: National Institute of Mental Health.

Sweeny, J., & Blaschke, C. (1980). How to analyze program costs: The Michigan study. *How to Evaluate Education Programs*, Ch. 34, 1-6.

Wade, R. K. (1984). What makes a difference in inservice teacher education: A meta-analysis of the research (Doctoral dissertation, University of Massachusetts, 1984). *Dissertation Abstracts International, 45*(1), 155A.

Wang, M. C., & Ellett, C. D. (August, 1981). *Program validation: The state of the art*. Unpublished manuscript, University of Pittsburgh, Learning Research and Development Center, Pittsburgh.

Weinrott, M. R., Jones, R. R., & Howard, J. R. (1980). How to evaluate cost effectiveness. *How to Evaluate Education Programs*, Ch. 65, 1-7.

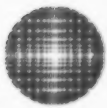
Weiss, C. H. (1972). *Evaluation research: Methods of assessing program effectiveness.* Englewood Cliffs, NJ: Prentice-Hall.

White, O. R. (1984). Selected issues in program evaluation: Arguments for the individual. *Advances in special education* (Vol 4. pp.69-121).

Williams, W., & Elmore, R. F. (Eds.). (1976). *Social program implementation*. New York: Academic Press.

Wood, F. H., & Thompson, S. R. (1980). Guidelines for better staff development. *Educational Leadership, 37*, 374-378.

Wood, F. H., Thompson, S. R., & Russell, F. Sr. (1981). Designing effective staff development programs. In B. Dillon-Peterson (Ed.), *Staff development/organization development* (pp.59-91). Alexandria, VA: Association for Supervision and Curriculum Development.



Washington
Research
Organization

Research in Education of
the Severely Handicapped

University of Washington
110 Miller Hall, DQ-05
Seattle, WA 98195